Journal of Computational and Applied Mathematics 44 (1992) 131–165
North-Holland

CAM 1201

Finite-element solution of flow problems with trailing conditions

Miloslav Feistauer, Jiří Felcman, Mirko Rokyta and Zdeněk Vlášek
Faculty of Mathematics and Physics, Charles University Prague, Czechoslovakia

Received 7 February 1990
Revised 5 December 1991

Abstract

Feistauer, M., J. Felcman, M. Rokyta and Z. Vlášek, Finite-element solution of flow problems with trailing conditions, Journal of Computational and Applied Mathematics 44 (1992) 131–165.

The paper deals with the finite-element solution of stream function problems describing nonviscous subsonic irrotational flows past profiles. The main emphasis is laid on the treatment of the nonstandard trailing stagnation conditions which lead to physically admissible solutions. The paper presents a general conception of stream function finite-element modelling of complicated flow problems and a complete theory of the finite-element approximations, including the investigation of the existence and uniqueness of the solution of the nonsymmetric discrete problem and the convergence of approximate solutions to the exact solution.

Keywords: Irrotational subsonic flow; stream function; trailing condition; stagnation point; finite elements; existence and uniqueness of approximate solutions; convergence; applications.

0. Introduction

In the study of real flows we use suitable mathematical models for obtaining necessary information about the structure of flow processes and aerodynamical properties of technological devices. Here we shall deal with the finite-element solution of stream function problems describing two-dimensional stationary irrotational nonviscous incompressible or subsonic compressible flows past profiles. These problems play an important role in the design of turbines, compressors and aircrafts.

Special attention will be paid to nonstandard problems with the so-called Kutta–Joukowskii trailing stagnation conditions which must be considered in order to get solutions admissible from the physical point of view. In the flow past a profile, which is smooth except a sharp trailing edge, we demand the velocity to be bounded. In case of a smooth profile we have concluded on the basis of experiments and calculations that for obtaining a physically reason-

Correspondence to: Prof. M. Feistauer, Faculty of Mathematics and Physics, Charles University Prague, Sokolovská 83, 18600 Praha 8, Czechoslovakia.

able solution it is sufficient to choose the trailing stagnation point, where the velocity is assumed to be zero, as a point on the backward part of the profile (with respect to the flow direction) with the greatest curvature.

In the sequel we introduce a general formulation which includes a series of flow problems of practical interest. The discretization via the finite-element method leads to the nonstandard, generally nonsymmetric discrete problem. We deal with its properties and prove the existence and uniqueness of the approximate finite-element solution and study the convergence of approximate solutions to the exact solution. We also describe a convergent iterative method for finding the solution of the discrete problem. The paper is concluded by examples of flows past profiles and cascades of profiles.

1. Continuous problem

By $\mathbb{R}^k$ we shall denote the k-dimensional Euclidean space. As a rule, we use the notation $x = (x_1, x_2)$ for points of the space $\mathbb{R}^2$. By $|\cdot|$ we shall denote the norm in $\mathbb{R}^k$. If $M \subset \mathbb{R}^k$, then $\overline{M}$ and ∂M denote the closure and the boundary of M, respectively. Let $\Omega \subset \mathbb{R}^2$ be a bounded domain. In the sequel we shall work with spaces of continuously differentiable functions $C^k(\overline{\Omega})$, $C^{k,\alpha}(\overline{\Omega})$ and the Lebesgue and Sobolev spaces $L^p(\Omega)$ and $W^{k,p}(\Omega)$, respectively. For their definitions and properties see, e.g., [4,20,22].

We shall start with two examples.

1.1. Plane or axially symmetric stationary irrotational nonviscous incompressible or subsonic compressible flow

This is (in a domain $\Omega \subset \mathbb{R}^2$) described by the equation

$$\sum_{i=1}^{2} \frac{\partial}{\partial x_i}\left(b(x, |\nabla u|^2)\frac{\partial u}{\partial x_i}\right) = 0, \quad \text{in } \Omega, \tag{1.1}$$

for the stream function u [6,8,15,18,23]. Here $b : \overline{\Omega} \times [0, +\infty) \to \mathbb{R}^1$ is a function given by the dependence of the density on the velocity. Its properties are specified in Section 1.3.

Let us assume that Ω is a bounded multiply-connected domain with the boundary $\partial\Omega$ formed by $r+1$, $r \geqslant 1$, disjoint simple closed curves C_0 (outer component) and $C_1, \ldots, C_r \subset \operatorname{Int} C_0$ (inner components) — see Fig. 1.1. Here $\operatorname{Int} C_0$ denotes the bounded component of $\mathbb{R}^2 \setminus C_0$. The curves $C_1, \ldots, C_r$ represent, e.g., profiles (airfoils) inserted into a bounded domain $\operatorname{Int} C_0$.

Let us specify boundary conditions on $\partial\Omega$ added to (1.1). On the outer part C_0 of $\partial\Omega$ it is suitable to consider mixed Dirichlet–Neumann conditions:

$$u \,|\, \Gamma_D = u_D, \qquad b\frac{\partial u}{\partial n} \,|\, \Gamma_N = -\phi_N. \tag{1.2}$$

$\partial/\partial n$ denotes the derivative with respect to the outer unit normal to $\partial\Omega$, u_D and ϕ_N are given functions, $C_0 = \overline{\Gamma}_D \cup \overline{\Gamma}_N$ and $\Gamma_D \cap \Gamma_N = \emptyset$. Let the profiles $C_1, \ldots, C_r$ be impermeable. Then the stream function is equal to a priori unknown constants q_i on C_i:

$$u \,|\, C_i = q_i, \quad i = 1, \ldots, r, \tag{1.3}$$

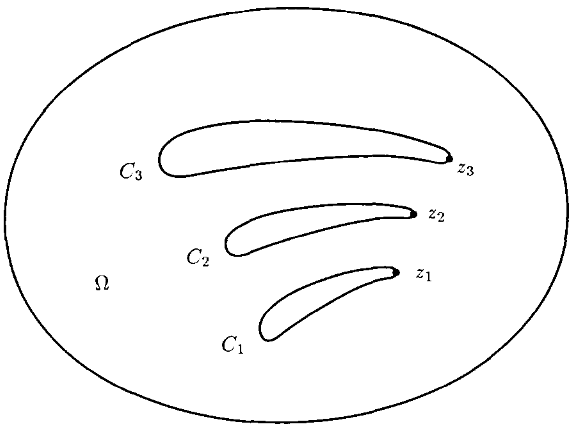

Fig. 1.1.

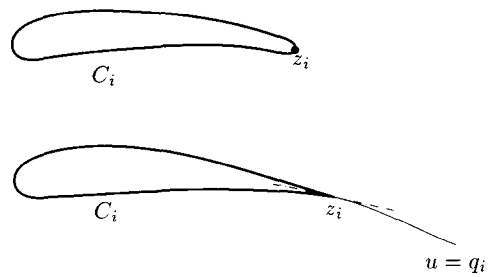

Fig. 1.2.

which must be determined in such a way that u satisfies suitable *complementary conditions*. Provided the *velocity circulation* γ_i around C_i is prescribed, then

$$\int_{C_i} b(\cdot, |\nabla u|^2)\, \mathrm{d}S = -\gamma_i. \tag{1.4}$$

If γ_i is unknown, it is necessary to choose q_i in such a way that the resulting solution is physically admissible, i.e., it satisfies the *Kutta–Joukowskii trailing condition*. If the profile C_i is smooth, then we consider the condition

$$\frac{\partial u}{\partial n}(z_i) = 0 \tag{1.5}$$

at a given *trailing stagnation point* $z_i \in C_i$. Provided the profile C_i is *smooth except exactly one sharp edge* at a point z_i, then we have to choose q_i in such a way that

$$|\nabla u|^2 \text{ is bounded in a neighbourhood of } C_i. \tag{1.6}$$

Using conformal mapping arguments [3] we find out that for the admissible plane irrotational incompressible flow satisfying (1.6) the following condition is valid.

Condition 1.1. *The stream line defined as a curve given by the condition $u = q_i$ leaves the profile C_i at the point z_i in the direction of the axes of the angle formed by the smooth parts of C_i.*

(See Fig. 1.2.)

In the sequel we shall also use (1.5) or Condition 1.1 in the study of a general fluid flow (compressible *or* axially symmetric *or* in a variable thickness fluid layer).

1.2. Flow past a cascade of profiles

The flow in a fluid layer bounded by two axially symmetric stream surfaces S_1, S_2 (close to each other) represents one of the most general two-dimensional models approximating three-

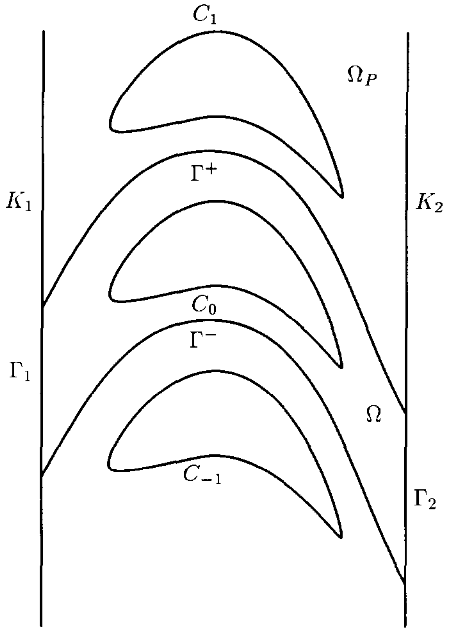

Fig. 1.3.

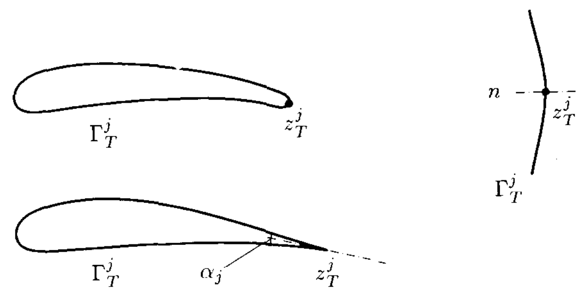

Fig. 1.4.

dimensional flows in cascades of blades of turbines, compressors and pumps. By introducing suitable coordinates x_1, x_2 on S_1, this surface and its intersections with the blades can be transformed into the plane (x_1, x_2), where we get a domain Ω_P. Its boundary $\partial\Omega_P$ is formed by two straight lines $K_i = \{(x_1, x_2);\ x_1 = d_i,\ x_2 \in \mathbb{R}^1\}$, $i = 1, 2$, $d_1 < d_2$, and by an infinite number of disjoint simple closed curves C_k, $k = 0, \pm 1, \pm 2, \ldots$, periodically spaced in the direction of x_2 with a period $\tau > 0$ (see Fig. 1.3). The curves C_k form the so-called *cascade of profiles*, the lines K_1 and K_2 represent the *inlet* and the *outlet* of the cascade, respectively.

The domain Ω_P is periodic in the direction of x_2 with the period τ. Assuming that the quantities describing the flow are also periodic in the direction of x_2 with the period τ, we can formulate the flow problem in a bounded domain Ω, which represents one period of Ω_P. That is, $\Omega \subset \Omega_P$ and $\partial\Omega$ consists of two components — the profile C_0 (inner component) and the union $\Gamma_1 \cup \Gamma_2 \cup \Gamma^- \cup \Gamma^+$ (outer component), where $\Gamma_i \subset K_i$ is a segment of the length τ, Γ^- is a piecewise linear arc and $\Gamma^+ = \{(x_1, x_2 + \tau);\ (x_1, x_2) \in \Gamma^-\}$. The initial points of Γ^-, Γ^+ lie on K_1, their terminal points lie on K_2 and all their other points are elements of Ω_P. See Fig. 1.3.

The quasi-stationary irrotational nonviscous incompressible or subsonic compressible flow on the axially symmetric stream surface S_1 in a layer of variable thickness past a cascade rotating with an angular velocity ω can be formulated as the following stream function problem in Ω [13,19]. Find $u : \overline{\Omega} \to \mathbb{R}_1$ and q_0, $q_1 \in \mathbb{R}^1$ satisfying the equation

$$\sum_{i=1}^{2} \frac{\partial}{\partial x_i}\left(b(x, |\nabla u|^2)\frac{\partial u}{\partial x_i}\right) = \omega\frac{\partial r^2}{\partial x_1}, \quad \text{in } \Omega, \tag{1.7}$$

and the conditions

$$u \mid \Gamma_1 = \Psi_1 + q_1, \tag{1.8a}$$

$$u \mid \Gamma_2 = \Psi_2, \tag{1.8b}$$

$$u \mid C_0 = q_0, \tag{1.9}$$

$$u(x_1, x_2 + \tau) = u(x_1, x_2) + Q, \quad (x_1, x_2) \in \Gamma^-, \tag{1.10a}$$

$$\left[b(\cdot, |\nabla u|^2) \frac{\partial u}{\partial n} \right] (x_1, x_2 + \tau) = -\left[b(\cdot, |\nabla u|^2) \frac{\partial u}{\partial n} \right] (x_1, x_2), \quad (x_1, x_2) \in \Gamma^-, \tag{1.10b}$$

$$\frac{1}{\tau} \int_{\Gamma_1} b(\cdot, |\nabla u|^2) \frac{\partial u}{\partial n} \, \mathrm{d}S = -\bar{\mu}_1 - \frac{\omega}{\tau} \int_{\Gamma_1} r^2 \, \mathrm{d}S \tag{1.11}$$

and

$$\frac{\partial u}{\partial n}(z_0) = 0, \tag{1.12}$$

provided the profile C_0 is smooth. Here the function b is given by the dependence of the density on the velocity. Moreover, both b and r depend on the geometry of the fluid layer. Ψ_1 and Ψ_2 are given functions, ω, Q, τ, $\bar{\mu}_1$ are given constants and $z_0 \in C_0$ is a given trailing stagnation point.

If the profile C_0 is smooth except a sharp edge at z_0, then (1.12) is replaced by (1.6) or Condition 1.1, where we set $i = 0$.

1.3. General problem

From the above examples we see that in problems of nonviscous fluid dynamics beside usual Dirichlet and Neumann boundary conditions we meet *nonstandard conditions*: periodicity conditions (1.10) and incomplete Dirichlet conditions (1.3), (1.8a), (1.9) combined with complementary integral conditions (1.4), (1.11) or trailing stagnation conditions (1.5), (1.12). On the basis of a detailed analysis of these examples [16,17,19], and also other types of flow problems, a *unified conception for the stream function finite-element modelling of nonviscous flows* was worked out.

The conception is based on the following *fundamental assumptions*.

(1) $\Omega \subset \mathbb{R}^2$ is a bounded domain with a piecewise smooth and Lipschitz-continuous boundary $\partial\Omega$ (cf. [4,20,22]).

(2) $$\partial\Omega = \bar{\Gamma}_{\mathrm{D}} \cup \bar{\Gamma}_{\mathrm{N}} \cup \left(\bigcup_{j=1}^{K_{\mathrm{I}}} \bar{\Gamma}_{\mathrm{I}}^j \right) \cup \left(\bigcup_{j=1}^{K_{\mathrm{T}}} \bar{\Gamma}_{\mathrm{T}}^j \right) \cup \bar{\Gamma}_{\mathrm{P}}^- \cup \bar{\Gamma}_{\mathrm{P}}^+,$$

where Γ_{I}^j are open arcs (i.e., arcs without their endpoints) or simple closed curves; Γ_{P}^-, Γ_{P}^+ are piecewise linear open arcs, $\bar{\Gamma}_{\mathrm{P}}^+ = Z_{\mathrm{P}}(\bar{\Gamma}_{\mathrm{P}}^-)$, where $Z_{\mathrm{P}} : \Gamma_{\mathrm{P}}^- \to^{\mathrm{onto}} \Gamma_{\mathrm{P}}^+$ is a continuous one-to-one mapping (usually representing translation or revolution); Γ_{D} and Γ_{N} are formed by a finite number of open arcs, $\Gamma_{\mathrm{D}} \neq \emptyset$. Γ_{T}^j is either a smooth open arc with a given point z_{T}^j or a simple closed smooth curve with a given trailing point $z^j \in \Gamma_{\mathrm{T}}^j$ or a simple closed curve smooth except exactly one point z_{T}^j at which the smooth parts of Γ_{T}^j form an inner angle $\alpha_j \in [0, \pi)$, see Fig. 1.4. All these arcs and curves are mutually disjoint.

(3) The *differential equation* has the form

$$\sum_{i=1}^{2} \frac{\partial}{\partial x_i}\left(b(\cdot, |\nabla u|^2)\frac{\partial u}{\partial x_i}\right) = \sum_{i=1}^{2} \frac{\partial f_i}{\partial x_i}, \quad \text{in } \Omega. \tag{1.13}$$

(4) We admit the following *boundary conditions*:

$$u \mid \Gamma_{\mathrm{D}} = u_{\mathrm{D}} \qquad \text{(Dirichlet condition (D))}, \tag{1.14}$$

$$b(\cdot, |\nabla u|^2)\frac{\partial u}{\partial n} \mid \Gamma_{\mathrm{N}} = \phi_{\mathrm{N}} \quad \text{(Neumann condition (N))}, \tag{1.15}$$

$$u(Z_{\mathrm{P}}(x)) = u(x) + Q, \tag{1.16a}$$

$$-\left[b(\cdot, |\nabla u|^2)\frac{\partial u}{\partial n}\right](Z_{\mathrm{P}}(x))$$

$$= \left[b(\cdot, |\nabla u|^2)\frac{\partial u}{\partial n}\right](x) - \sum_{i=1}^{2} (f_i n_i)(Z_{\mathrm{P}}(x)) - \sum_{i=1}^{2} (f_i n_i)(x), \quad x \in \Gamma_{\mathrm{P}}^-$$

$$\text{(periodicity conditions (P))}, \tag{1.16b}$$

$$u \mid \Gamma_{\mathrm{I}}^j = u_{\mathrm{I}}^j + q_{\mathrm{I}}^j, \quad q_{\mathrm{I}}^j = \text{const.}, \tag{1.17a}$$

$$\int_{\Gamma_{\mathrm{I}}^j} b(\cdot, |\nabla u|^2)\frac{\partial u}{\partial n}\,\mathrm{d}S = -\gamma_{\mathrm{I}}^j + \sum_{i=1}^{2}\int_{\Gamma_{\mathrm{I}}^j} f_i n_i\,\mathrm{d}S, \quad j = 1, \ldots, K_{\mathrm{I}}$$

$$\big(\text{incomplete Dirichlet conditions combined with integral conditions } (\mathrm{I}_j)\big), \tag{1.17b}$$

$$u \mid \Gamma_{\mathrm{T}}^j = u_{\mathrm{T}}^j + q_{\mathrm{T}}^j, \quad q_{\mathrm{T}}^j = \text{const.}, \tag{1.18a}$$

$$\frac{\partial u}{\partial n}\left(z_{\mathrm{T}}^j\right) = v_{\mathrm{T}}^j, \quad \text{if } \Gamma_{\mathrm{T}}^j \text{ is smooth}, \tag{1.18b}$$

or

the streamline $u = q_{\mathrm{T}}^j$ leaves Γ_{T}^j in the direction of the axis of the angle α_j at the point z_{T}^j, if Γ_{T}^j is an impermeable profile with a sharp edge at z_{T}^j, $j = 1, \ldots, K_{\mathrm{T}}$, (incomplete Dirichlet conditions combined with trailing conditions (T_j)). (1.18b*)

u_{D}, ϕ_{N}, u_{I}^j, u_{T}^j are given continuous functions, Q, γ_{I}^j, v_{T}^j are given constants, $z_{\mathrm{T}}^j \in \Gamma_{\mathrm{T}}^j$ are given points; $u : \overline{\Omega} \to \mathbb{R}^1$ is an unknown function, q_{I}^j, $q_{\mathrm{T}}^j \in \mathbb{R}^1$ are unknown constants.

If Γ_{T}^j is an impermeable profile represented by a simple closed curve, we set $v_{\mathrm{T}}^j = 0$, $u_{\mathrm{T}}^j = 0$.

(5) The functions b, f_1, f_2 have the following properties. (a) $\overline{\Omega} \subset \tilde{\Omega} \subset \mathbb{R}^2$, $\tilde{\Omega}$ is an open set; the function b and its derivatives $\partial b/\partial\eta$, $\partial b/\partial x_i$, $i = 1, 2$, are continuous in $\tilde{\Omega} \times [0, +\infty)$; f_i, $\partial f_i/\partial x_j$, $i, j = 1, 2$, are continuous in $\tilde{\Omega}$.

(b) There exist constants $c_1, c_2, c_3, c_4 > 0$ such that

$$c_1 \leqslant b \leqslant c_2, \quad \text{in } \tilde{\Omega} \times [0, +\infty), \tag{1.19}$$

$$0 \leqslant \frac{\partial b}{\partial \eta} \leqslant c_3, \quad \left|\frac{\partial b}{\partial x_i}\right| \leqslant c_3, \; i = 1, 2, \qquad \text{in } \tilde{\Omega} \times [0, +\infty), \tag{1.20}$$

$$\left|\xi \frac{\partial b}{\partial \eta}(x, \xi^2)\right|, \left|\xi^2 \frac{\partial b}{\partial \eta}(x, \xi^2)\right| \leqslant c_4, \quad x \in \tilde{\Omega}, \; \xi \in \mathbb{R}^1. \tag{1.21}$$

If the flow is incompressible, then $b = b(x)$ and (1.13) is linear. In the compressible case this equation is nonlinear.

The contact of some boundary conditions is prohibited. In Table 1.1 the sign X denotes the possible contacts of parts of $\partial\Omega$ with different boundary conditions, the empty square denotes the prohibition.

In case of the contact allowed of the conditions (D)–(P) or (I_j)–(P) or (T_j)–(P) it is necessary to fulfil their *consistency*:

$$u_D(Z_P(x)) - u_D(x) = Q, \quad \text{if } x \in \bar{\Gamma}_P^- \cap \bar{\Gamma}_D \text{ and } Z_P(x) \in \bar{\Gamma}_P^+ \cap \bar{\Gamma}_D, \tag{1.22a}$$

$$u_I^j(Z_P(x)) - u_I^j(x) = Q, \quad \text{if } x \in \bar{\Gamma}_P^- \cap \bar{\Gamma}_I^j \text{ and } Z_P(x) \in \bar{\Gamma}_P^+ \cap \bar{\Gamma}_I^j, \tag{1.22b}$$

$$u_T^j(Z_P(x)) - u_T^j(x) = Q, \quad \text{if } x \in \bar{\Gamma}_P^- \cap \bar{\Gamma}_T^j \text{ and } Z_P(x) \in \bar{\Gamma}_P^+ \cap \bar{\Gamma}_T^j. \tag{1.22c}$$

If some of the conditions (1.14)–(1.18) are not considered, then the corresponding part of $\partial\Omega$ is assumed to be empty. Moreover, if we do not consider conditions (1.17) or (1.18), we set $K_I = 0$ or $K_T = 0$, respectively.

It is easy to find out that the examples of Sections 1.1 and 1.2 can be formulated in the frame of the general problem (1.13)–(1.18).

1.4. Classical solution

The solution of the above general problem is defined as a function $u \in C^2(\bar{\Omega})$ and constants q_I^j, $j = 1, \ldots, K_I$, q_T^j, $j = 1, \ldots, K_T$, satisfying (1.13)–(1.18).

Table 1.1

	D	N	P	I_1	I_{K_I}	T_1	T_{K_T}
D		X	X				
N	X		X				
P	X	X		X	X	X	X
I_1			X				
I_{K_I}			X				
T_1			X				
T_{K_T}			X				

2. Variational formulation of the problem

Let us assume that u is a classical solution of the problem (1.13)–(1.18). If we multiply (1.13) by an arbitrary $v \in C^1(\overline{\Omega})$, integrate over Ω and use Green's theorem in a standard way, we get the identity

$$\int_{\partial\Omega} b \frac{\partial u}{\partial n} v \,\mathrm{d}S - \int_{\Omega} b \,\nabla u \cdot \nabla v \,\mathrm{d}x = \int_{\partial\Omega} v \sum_{i=1}^{2} f_i n_i \,\mathrm{d}S - \int_{\Omega} \sum_{i=1}^{2} f_i \frac{\partial v}{\partial x_i} \,\mathrm{d}x, \tag{2.1}$$

which will be the basis for the variational formulation of the problem.

First, let us consider the following problem.

2.1. *Problem without trailing conditions*

When $K_{\mathrm{T}} = 0$, we can define a weak solution in the well-known Sobolev space $H^1(\Omega) = W^{1,2}(\Omega)$. Let

$$\begin{aligned} \nu = \{ v \in C^1(\overline{\Omega});\ & v \mid \Gamma_{\mathrm{D}} = 0,\ v \mid \Gamma_{\mathrm{I}}^j = \text{const.},\ j = 1, \ldots, K_{\mathrm{I}}, \\ & v(Z_{\mathrm{P}}(x)) = v(x),\ x \in \Gamma_{\mathrm{P}}^- \}. \end{aligned} \tag{2.2}$$

If we consider $v \in \nu$ in (2.1) and take into account the conditions (1.15), (1.16b), (1.17b), we get the relation

$$\begin{aligned} \int_{\Omega} b(\cdot, |\nabla u|^2)\, \nabla u \cdot \nabla v \,\mathrm{d}x = & - \int_{\Gamma_{\mathrm{N}}} \phi_{\mathrm{N}} v \,\mathrm{d}S - \sum_{j=1}^{K_{\mathrm{I}}} \gamma_{\mathrm{I}}^j v \mid \Gamma_{\mathrm{I}}^j \\ & + \int_{\Omega} \sum_{i=1}^{2} f_i \frac{\partial v}{\partial x_i} \,\mathrm{d}x, \quad v \in \nu. \end{aligned} \tag{2.3}$$

Further, we define the space

$$\begin{aligned} V = \{ v \in H^1(\Omega);\ & v \mid \Gamma_{\mathrm{D}} = 0,\ v \mid \Gamma_{\mathrm{I}}^j = \text{const.},\ j = 1, \ldots, K_{\mathrm{I}}, \\ & v(Z_{\mathrm{P}}(x)) = v(x),\ x \in \Gamma_{\mathrm{P}}^- \} \end{aligned} \tag{2.4}$$

(where the restrictions on Γ_{D}, Γ_{I}^j and Γ_{P}^- are meant in the sense of traces).

Let $u^* \in H^1(\Omega)$ satisfy the conditions

$$u^* \mid \Gamma_{\mathrm{D}} = u_{\mathrm{D}}, \tag{2.5a}$$

$$u^* \mid \Gamma_{\mathrm{I}}^j = u_{\mathrm{I}}^j, \quad j = 1, \ldots, K_{\mathrm{I}}, \tag{2.5b}$$

$$u^*(Z_{\mathrm{P}}(x)) = u^*(x) + Q, \quad x \in \Gamma_{\mathrm{P}}^-. \tag{2.5c}$$

We denote

$$\begin{aligned} a(u, v) &= \int_{\Omega} b(\cdot, |\nabla u|^2)\, \nabla u \cdot \nabla v \,\mathrm{d}x, \quad u, v \in H^1(\Omega), \\ L(v) &= - \int_{\Gamma_{\mathrm{N}}} \phi_{\mathrm{N}} v \,\mathrm{d}S - \sum_{j=1}^{K_{\mathrm{I}}} \gamma_{\mathrm{I}}^j v \mid \Gamma_{\mathrm{I}}^j + \int_{\Omega} \sum_{i=1}^{2} f_i \frac{\partial v}{\partial x_i} \,\mathrm{d}x, \quad v \in V. \end{aligned} \tag{2.6}$$

Definition 2.1. We say that $u:\bar{\Omega} \to \mathbb{R}^1$ is a *weak solution* of problem (1.13)–(1.17) (without trailing conditions), if

$$u \in H^1(\Omega), \tag{2.7a}$$

$$u - u^* \in V, \tag{2.7b}$$

$$a(u, v) = L(v), \quad \forall v \in V. \tag{2.7c}$$

Provided the closure of ν in $H^1(\Omega)$ is V (cf., e.g., [13]), problems (1.13)–(1.18) and (2.7) are formally equivalent. The existence and uniqueness of the weak solution is a consequence of the properties of the functions b, f_1, f_2 and the monotone operator theory [13,19,23].

Moreover, in [13] a *convergent iterative process for the solution of problem* (2.7) was proposed. Let us put $\tilde{b}(x) = b(x, 0)$ and define the form

$$(u, v)_{\tilde{b}} = \int_\Omega \tilde{b}\, \nabla u \cdot \nabla v \,\mathrm{d}x, \quad u, v \in H^1(\Omega), \tag{2.8}$$

which is a scalar product on V in virtue of the properties of b. Let $\{u^k\}$ be the sequence defined in the following way:

$$u^0 \in H^1(\Omega), \qquad u^0 - u^* \in V, \qquad (u^0, v)_{\tilde{b}} = L(v), \quad \forall v \in V, \tag{2.9a}$$

$$\begin{aligned} &u^{k+1} \in H^1(\Omega), \qquad u^{k+1} - u^* \in V,\\ &(u^{k+1}, v)_{\tilde{b}} = (u^k, v)_{\tilde{b}} - \nu\big(a(u^k, v) - L(v)\big), \quad v \in V, \end{aligned} \tag{2.9b}$$

for $k \geqslant 0$.
There exists $\bar{\nu} > 0$ such that for each $\nu \in (0, \bar{\nu})$ the sequence $\{u^k\}$ converges to the solution u of (2.7).

It is possible to find out that (2.9) is a *steepest descent* method with preconditioning defined by the form $(\cdot, \cdot)_{\tilde{b}}$ for minimizing the functional

$$\Phi(u) = \tfrac{1}{2}\int_\Omega \left(\int_0^{|\nabla u|^2} b(x, \eta)\,\mathrm{d}\eta \right) \mathrm{d}x - L(u), \tag{2.10}$$

on $u^* + V$.

2.2. "Variational" formulation of the problem with trailing conditions

Because of the discrete trailing conditions (1.18b) or (1.18b*), the concept of the weak solution $u \in H^1(\Omega)$ has no sense and therefore we shall work with classical solutions. Let us consider condition (1.18b). (The case of (1.18b*) will be treated later in the discretization process.)

Let us assume the existence of a function $u^* \in C^2(\bar{\Omega})$ such that

$$u^* \mid \Gamma_\mathrm{D} = u_\mathrm{D}, \tag{2.11a}$$

$$u^* \mid \Gamma_\mathrm{I}^j = u_\mathrm{I}^j, \qquad j = 1, \ldots, K_\mathrm{I}, \tag{2.11b}$$

$$u^* \mid \Gamma_\mathrm{T}^j = u_\mathrm{T}^j, \qquad j = 1, \ldots, K_\mathrm{T}, \tag{2.11c}$$

$$\frac{\partial u^*}{\partial n}(z_j) = v_\mathrm{T}^j, \qquad j = 1, \ldots, K_\mathrm{T}, \tag{2.11d}$$

$$u^*(Z_\mathrm{P}(x)) = u^*(x) + Q, \quad x \in \Gamma_\mathrm{P}^-. \tag{2.11e}$$

Further, we set

$$V=\{v \in C^1(\Omega);\ v \mid \Gamma_{\mathrm{D}}=0,\ v \mid \Gamma_{\mathrm{T}}^j=0,\ j=1,\dots,K_{\mathrm{T}},$$
$$v \mid \Gamma_{\mathrm{I}}^j=\text{const.},\ j=1,\dots,K_{\mathrm{I}},\ v(Z_{\mathrm{P}}(x))=v(x),\ x \in \Gamma_{\mathrm{P}}^-\}, \tag{2.12}$$

$$\tilde{V}=\Big\{v \in C^2(\Omega);\ v \mid \Gamma_{\mathrm{D}}=0,\ v \mid \Gamma_{\mathrm{T}}^j=\text{const.},\ \frac{\partial v}{\partial n}(z_{\mathrm{T}}^j)=0,\ j=1,\dots,K_{\mathrm{T}},$$
$$v \mid \Gamma_{\mathrm{I}}^j=\text{const.},\ j=1,\dots,K_{\mathrm{I}},\ v(Z_{\mathrm{P}}(x))=v(x),\ x \in \Gamma_{\mathrm{P}}\Big\}. \tag{2.13}$$

Now, on the basis of (2.1) we find out that $u:\bar{\Omega} \to \mathbb{R}^1$ is a classical solution of problem (1.13)–(1.18) if and only if

$$u \in C^2(\bar{\Omega}), \tag{2.14a}$$

$$u-u^* \in \tilde{V}, \tag{2.14b}$$

$$a(u,v)=L(v), \quad \forall v \in V, \tag{2.14c}$$

where the forms a and L are defined in (2.6).

The solvability of problems with trailing condition (1.18b) was studied for special cases in [7,10,12,15]. The solvability of the flow past an isolated profile with condition (1.6) was proved in [6].

3. Finite-element discretization of the problem

Let us approximate Ω by a domain Ω_h with a piecewise linear boundary $\partial\Omega_h$, the vertices of which lie on $\partial\Omega$. Let $\mathscr{T}_h$ be a triangulation of Ω_h, i.e., a set consisting of a finite number of closed triangles, which has the following properties.

Property 3.1. (a) $\bar{\Omega}_h=\bigcup_{T \in \mathscr{T}_h} T$;

(b) if $T_1, T_2 \in \mathscr{T}_h$, $T_1 \neq T_2$, then either $T_1 \cap T_2=\emptyset$ or $T_1 \cap T_2$ is a common vertex or $T_1 \cap T_2$ is a common side of T_1, T_2.

We denote by $\sigma_h=\{P_1,\dots,P_N\}$ the set of all vertices of $\mathscr{T}_h$ and assume that the next property holds.

Property 3.2. (a) $\sigma_h \subset \bar{\Omega}_h$, $\sigma_h \cap \partial\Omega_h \subset \partial\Omega$;

(b) the intersections of the parts of $\partial\Omega$ with different types of boundary conditions are elements of σ_h;

(c) the points where $\partial\Omega$ is not smooth are elements of σ_h;

(d) provided $\Gamma_{\mathrm{P}}^- \neq \emptyset$, we suppose that $\bar{\Gamma}_{\mathrm{P}}^- \cup \bar{\Gamma}_{\mathrm{P}}^+ \subset \partial\Omega_h$ and $P_j \in \bar{\Gamma}_{\mathrm{P}}^- \cap \sigma_h \Leftrightarrow Z_{\mathrm{P}}(P_j) \in \bar{\Gamma}_{\mathrm{P}}^+ \cap \sigma_h$;

(e) $z_{\mathrm{T}}^j \in \sigma_h$, $j=1,\dots,K_{\mathrm{T}}$; to each z_{T}^j there exists a triangle $T_j \in \mathscr{T}_h$ with vertices $\tilde{P}_j=z_{\mathrm{T}}^j$ and $P_j^* \in \Omega_h$ such that the side $S_j=\tilde{P}_j P_j^* \subset \partial T_j$ is normal to Γ_{T}^j, if Γ_{T}^j is smooth, and S_j has the direction of the axis of the angle at z_{T}^j, if Γ_{T}^j is a profile smooth except the point z_{T}^j (cf. Section 1.3);

(f) ϕ_{N} is piecewise smooth and the points where ϕ_N is not smooth belong to σ_h.

The sets Γ_{D}, Γ_{N}, Γ_{I}^j, $\Gamma_{\mathrm{T}}^j \subset \partial\Omega$ are approximated in a natural way by piecewise linear arcs and curves $\Gamma_{\mathrm{D}h}$, $\Gamma_{\mathrm{N}h}$, $\Gamma_{\mathrm{I}h}^j$, $\Gamma_{\mathrm{T}h}^j \subset \partial\Omega_h$.

By the symbols h and θ_h we shall denote the length of the maximum side and the magnitude of the smallest angle, respectively, in the triangulation $\mathscr{T}_h$. We assume that h is sufficiently small, so that $\overline{\Omega}_h \subset \tilde{\Omega}$.

The approximate solution of the problem will be sought in the space of *conforming linear triangular elements*

$$X_h = \left\{v_h \in C\left(\overline{\Omega}_h\right);\ v_h \mid T \text{ is linear } \forall T \in \mathscr{T}_h\right\}, \tag{3.1}$$

and the space V of test functions defined in (2.12) will be approximated by

$$V_h = \Big\{v_h \in X_h;\ v_h \mid \Gamma_{\mathrm{D}h} \cup \left(\bigcup_{j=1}^{K_{\mathrm{T}}} \Gamma_{\mathrm{T}h}^j\right) = 0,\ v_h \mid \Gamma_{\mathrm{I}h}^j = \text{const.},\ j = 1,\dots,K_{\mathrm{I}},$$

$$v_h\left(Z_{\mathrm{P}}(P_j)\right) = v_h(P_j),\ P_j \in \overline{\Gamma}_{\mathrm{P}}^- \cap \sigma_h\Big\}. \tag{3.2}$$

Let us notice that $\nabla v_h \mid T = \text{const.}$ for each $v_h \in X_h$ and $T \in \mathscr{T}_h$.

In the discretization of condition (1.18b) we shall use the finite-difference approach proposed in [11]. Provided Γ_{T}^j is smooth and $u \in C^2(\overline{\Omega})$ is an exact solution of problem (1.13)–(1.18), then

$$\frac{\partial u}{\partial n}\left(z_{\mathrm{T}}^j\right) = \frac{\partial u}{\partial n}\left(\tilde{P}_j\right) = \frac{u\left(\tilde{P}_j\right) - u\left(P_j^*\right)}{\left|\tilde{P}_j - P_j^*\right|} + \mathrm{O}\left(\left|\tilde{P}_j - P_j^*\right|\right), \tag{3.3}$$

which together with (1.18a) leads to the relation

$$u(x) = u\left(P_j^*\right) + u_{\mathrm{T}}^j(x) - u_{\mathrm{T}}^j\left(\tilde{P}_j\right)\left|\tilde{P}_j - P_j^*\right| v_{\mathrm{T}}^j + \mathrm{O}(h^2), \quad x \in \overline{\Gamma}_{\mathrm{T}}^j. \tag{3.4}$$

From this we come to the discretization of conditions (1.18a), (1.18b):

$$u_h(P_i) = q_{\mathrm{T}}^j = u_h\left(P_j^*\right) + u_{\mathrm{T}}^j(P_i) - u_{\mathrm{T}}^j\left(\tilde{P}_j\right)\left|\tilde{P}_j - P_j^*\right| v_{\mathrm{T}}^j, \quad P_i \in \sigma_h \cap \overline{\Gamma}_{\mathrm{T}}^j. \tag{3.5}$$

If $v_{\mathrm{T}}^j = 0$ and Γ_{T}^j is a smooth profile, then (3.5) represents a discrete form of the trailing condition:

$$u_h(P_i) = u_h\left(P_j^*\right), \quad P_i \in \sigma_h \cap \overline{\Gamma}_{\mathrm{T}}^j. \tag{3.5*}$$

If the profile Γ_{T}^j is smooth except the sharp trailing edge z_{T}^j, we shall use condition (1.18b*) which obviously leads to (3.5*). It means that we again have (3.5), where we set $v_{\mathrm{T}}^j = 0$, $u_{\mathrm{T}}^j = 0$.

Now it is suitable to introduce a function $u_h^* \in X_h$ with the properties

$$u_h^*(P_i) = u_{\mathrm{D}}(P_i), \qquad P_i \in \sigma_h \cap \overline{\Gamma}_{\mathrm{D}}, \tag{3.6a}$$

$$u_h^*(P_i) = u_{\mathrm{I}}^j(P_i), \qquad P_i \in \sigma_h \cap \overline{\Gamma}_{\mathrm{I}}^j, \quad j = 1,\dots,K_{\mathrm{I}}, \tag{3.6b}$$

$$u_h^*(P_i) = u_{\mathrm{T}}^j(P_i) - u_{\mathrm{T}}^j\left(\tilde{P}_j\right) + \left|\tilde{P}_j - P_j^*\right| v_{\mathrm{T}}^j, \quad P_i \in \sigma_h \cap \overline{\Gamma}_{\mathrm{T}}^j, \quad j = 1,\dots,K_{\mathrm{T}}, \tag{3.6c}$$

$$u_h^*(Z_{\mathrm{P}}(P_i)) = u_h^*(P_i) + Q, \qquad P_i \in \sigma_h \cap \overline{\Gamma}_{\mathrm{P}}^-, \tag{3.6d}$$

$$u_h^*\left(P_j^*\right) = 0, \qquad j = 1,\dots,K_{\mathrm{T}}, \tag{3.6e}$$

and seek for an approximate solution in the set $u_h^* + \tilde{V}_h$, where

$$\tilde{V}_h = \{v_h \in X_h;\ v_h \mid \Gamma_{\mathrm{D}h} = 0,\ v_h \mid \left(\Gamma_{\mathrm{T}h}^j \cup S_j\right) = \text{const. for } j = 1, \ldots, K_{\mathrm{T}},$$
$$v_h \mid \Gamma_{\mathrm{I}}^j = \text{const. for } j = 1, \ldots, K_{\mathrm{I}},\ v_h(Z_{\mathrm{P}}(x)) = v_h(x)$$
$$\text{for } x \in \Gamma_{\mathrm{P}}^-\}. \tag{3.7}$$

From (3.6c) and (3.6e) we set that $(\partial u_h^*/\partial n)(z_{\mathrm{T}}^j) = v_{\mathrm{T}}^j$ and $u_h^*(P_i) = u_{\mathrm{T}}^j(P_i) + \text{const.}$ for $P_i \in \sigma_h \cap \bar{\Gamma}_{\mathrm{T}}^j$. If $v_h \in \tilde{V}_h$, then $(\partial v_h/\partial n)(z_{\mathrm{T}}^j) = 0$.

Further, let us approximate the forms a and L by

$$\tilde{a}_h(u_h, v_h) = \int_{\Omega_h} b\left(\cdot, |\nabla u_h|^2\right) \nabla u_h \cdot \nabla v_h \, \mathrm{d}x, \quad u_h, v_h \in X_h, \tag{3.8}$$

and

$$\tilde{L}_h(v_h) = -\int_{\Gamma_{\mathrm{N}h}} \phi_{\mathrm{N}h} v_h \, \mathrm{d}S - \sum_{j=1}^{K_{\mathrm{I}}} \gamma_{\mathrm{I}}^j v_h \mid \Gamma_{\mathrm{I}h}^j + \int_{\Omega_h} \sum_{i=1}^{2} f_i \frac{\partial v_h}{\partial x_i} \, \mathrm{d}x, \quad v_h \in V_h, \tag{3.9}$$

respectively. Here $\phi_{\mathrm{N}h} : \bar{\Gamma}_{\mathrm{N}h} \to \mathbb{R}^1$ is a suitable approximation of ϕ_{N}.

In practical calculations the integrals in (3.8), (3.9) must be evaluated by quadrature formulae. We write

$$\int_{\Omega_h} = \sum_{T \in \mathscr{T}_h} \int_T, \qquad \int_T g \, \mathrm{d}x \approx \operatorname{meas}(T) \sum_{k=1}^{k_T} \alpha_{T,k} g(x_{T,k}), \quad \text{for } g \in C(T), \tag{3.10}$$

where meas(T) is the measure of T, $\alpha_{T,k} \in \mathbb{R}^1$ and $x_{T,k} \in T$. Let us assume that

$$\alpha_{T,k} \geqslant 0, \quad \sum_{k=1}^{k_T} \alpha_{T,k} = 1, \quad \forall T \in \mathscr{T}_h. \tag{3.11}$$

Similarly,

$$\begin{aligned} &\int_{\Gamma_{\mathrm{N}h}} \phi \, \mathrm{d}s = \sum_{S_r^h \subset \bar{\Gamma}_{\mathrm{N}h}} \int_{S_r^h} \phi \, \mathrm{d}s, \\ &\int_{S_r^h} \phi \, \mathrm{d}s \approx s_r^h \sum_{j=1}^{k_r} \beta_{r,j} \phi(x_{r,j}), \quad \text{for } \phi \in C(S_r^h). \end{aligned} \tag{3.12}$$

We denote by $S_r^h \subset \partial\Omega_h$ sides of triangles $T \in \mathscr{T}_h$ adjacent to $\partial\Omega$, s_r^h the length of S_r^h and suppose that $\beta_{r,j} \in \mathbb{R}^1$, $x_{r,j} \in S_r^h$.

If we choose $x_{r,j} \in \sigma_h \cap \bar{\Gamma}_{\mathrm{N}}$ and set $\phi_{\mathrm{N}h}(x_{r,j}) = \phi_{\mathrm{N}}(x_{r,j})$, it is not necessary to introduce the approximation $\phi_{\mathrm{N}h}$ of ϕ_{N}.

With the use of numerical integration, the forms $\tilde{a}_h$ and $\tilde{L}_h$ are approximated by

$$a_h(u_h, v_h) = \sum_{T \in \mathscr{T}_h} \mathscr{I}_T(u_h) \nabla u_h \mid T \cdot \nabla v_h \mid T, \quad u_h, v_h \in X_h, \tag{3.13}$$

where

$$\mathscr{I}_T(u_h) = \operatorname{meas}(T) \sum_{k=1}^{k_T} \alpha_{T,k} b\left(x_{T,k}, |\nabla u_h \mid T|^2\right) \approx \int_T b\left(\cdot, |\nabla u_h|^2\right) \mathrm{d}x, \tag{3.14}$$

and

$$L_h(v_h) = -\sum_{S_r^h \subset \bar{\Gamma}_{Nh}} s_r^h \sum_{j=1}^{k_r} \beta_{r,j} \phi_{Nh}(x_{r,j}) v(x_{r,j}) - \sum_{j=1}^{K_I} \gamma_I^j v_h \mid \Gamma_{Ih}^j$$

$$+ \sum_{T \in \mathscr{T}_h} \operatorname{meas}(T) \sum_{j=1}^{k_T} \alpha_{T,j} \sum_{i=1}^{2} f_i(x_{T,j}) \frac{\partial v}{\partial x_i} \mid T, \quad v_h \in V_h. \tag{3.15}$$

By (1.19),

$$c_1 \leqslant \mathscr{I}_T(u_h) \operatorname{meas}^{-1}(T) \leqslant c_2, \quad u_h \in X_h, \ T \in \mathscr{T}_h. \tag{3.16}$$

The forms $a_h(u_h, v_h)$ and $L_h(v_h)$ are *linear with respect to* v_h.

3.1. Discrete problem

Find $u_h : \bar{\Omega}_h \to \mathbb{R}^1$ such that

$$u_h \in X_h, \tag{3.17a}$$

$$u_h - u_h^* \in \tilde{V}_h, \tag{3.17b}$$

$$a_h(u_h, v_h) = L_h(V_h), \quad \forall v_h \in V_h. \tag{3.17c}$$

u_h is called an *approximate solution*.

Let us notice that for $K_T = 0$ we have $\tilde{V}_h = V_h$. However, if $K_T > 0$, then $\tilde{V}_h \neq V_h$ and the discrete problem reminds the Petrov–Galerkin method (see, e.g., [21]. However, the structure of problem (3.17) is different from problems treated in [21].

4. The study of the discrete problem

4.1. Properties of the spaces V_h and $\tilde{V}_h$

In the space X_h there exists a basis formed by functions w_i, $i = 1, \ldots, N$, such that

$$w_i(P_j) = \delta_{ij} \quad (\text{Kronecker symbol}), \ i, j = 1, \ldots, N. \tag{4.1}$$

If $v_h \in X_h$, then

$$v_h = \sum_{j=1}^{N} v_h(P_j) w_j. \tag{4.2}$$

In V_h there exists a basis $\{w_i^*\}_{i=1}^n$,

$$n = \operatorname{card}\left\{\sigma_h \cap \left[\left(\Omega_h \cup \bar{\Gamma}_{Nh} \cup \bar{\Gamma}_{Ph}^-\right) \setminus \left(\bar{\Gamma}_{Dh} \cup \left(\bigcup_{j=1}^{K_I} \bar{\Gamma}_{Ih}^j\right) \cup \left(\bigcup_{j=1}^{K_T} \bar{\Gamma}_{Th}^j\right)\right)\right]\right\} + K_I,$$

formed by the functions (provided the points $P_i \in \sigma_h$ are numbered in a suitable way)

$$w_i \text{ for } P_i \in \sigma^{(1)} := \sigma_h \cap \left[\Omega_h \cup \bar{\Gamma}_{Nh} \setminus \left(\bar{\Gamma}_{Dh} \cup \bar{\Gamma}^-_{Ph} \cup \bar{\Gamma}^+_{Ph} \cup \{P_j^*, j = 1, \dots, K_T\}\right)\right], \tag{4.3a}$$

$$w_i + w_j \text{ for } P_i \in \sigma^{(2)} := \sigma_h \cap \left[\bar{\Gamma}^-_{Ph} \setminus \left(\bar{\Gamma}_{Dh} \cup \left(\bigcup_{j=1}^{K_I} \bar{\Gamma}^j_{Ih}\right) \cup \left(\bigcup_{j=1}^{K_T} \bar{\Gamma}^j_{Th}\right)\right)\right], \quad P_j = Z_P(P_i), \tag{4.3b}$$

$$\sum_{P_j \in \sigma_h \cap \bar{\Gamma}^i_I} w_j, \quad i = 1, \dots, K_I, \tag{4.3c}$$

$$w_i \text{ for } P_i = P_j^*, \quad j = 1, \dots, K_T. \tag{4.3d}$$

In $\tilde{V}_h$ we have the basis $\{\tilde{w}_i^*\}_{i=1}^n$ formed by the functions (4.3a)–(4.3c) and

$$\sum_{P_i \in \sigma_h \cap \bar{\Gamma}^j_T \cup \{P_j^*\}} w_i, \quad j = 1, \dots, K_T. \tag{4.3$\tilde{\text{d}}$}$$

It is evident that there exists a one-to-one correspondence between basis functions w_i^* and $\tilde{w}_i^*$, since the functions w_i^* and $\tilde{w}_i^*$ of the types (4.3a)–(4.3c) are identical and the basis functions of the types (4.3d) and (4.3$\tilde{\text{d}}$) are associated with the points $P_j^* \in \sigma_h \cap \Omega_h$, $j = 1, \dots, K_T$. Hence,

$$\tilde{w}_i^* \geqslant w_i^*, \quad i = 1, \dots, n. \tag{4.4}$$

Proof of the above assertions. This follows from (4.2), definitions (3.2) and (3.7) of the spaces V_h and $\tilde{V}_h$ and the fact that in view of Table 1.1 the set

$$\sigma_h \setminus \left(\bar{\Gamma}_{Dh} \cup \left(\bigcup_{j=1}^{K_T} \bar{\Gamma}^j_{Th}\right)\right)$$

is a sum of disjoint sets

$$\begin{aligned} &\{P_i\}, \quad P_i \in \sigma^{(1)}, \qquad \{P_i, P_j\}, \quad P_i \in \sigma^{(2)}, \; P_j = Z_P(P_i), \\ &\left\{P_j;\; P_j \in \sigma_h \cap \bar{\Gamma}^i_{Ih}\right\}, \quad i = 1, \dots, K_I, \\ &\left\{P_i = P_j^*\right\}, \quad j = 1, \dots, K_T, \end{aligned} \tag{4.5}$$

and $\sigma_h \setminus \bar{\Gamma}_{Dh}$ is a sum of disjoint sets

$$\begin{aligned} &\{P_i\}, \quad P_i \in \sigma^{(1)}, \qquad \{P_i, P_j\}, \quad P_i \in \sigma^{(2)}, \; P_j = Z_P(P_i), \\ &\left\{P_j;\; P_j \in \sigma_h \cap \bar{\Gamma}^i_{Ih}\right\}, \quad i = 1, \dots, K_I, \\ &\left\{P_j;\; P_j \in \sigma_h \cap \left(\bar{\Gamma}^i_{Th} \cup \{P_i^*\}\right)\right\}, \quad i = 1, \dots, K_T. \qquad \square \end{aligned} \tag{4.5*}$$

If $z_h \in \tilde{V}_h$, then

$$z_h = \sum_{i=1}^n z_i \tilde{w}_i^*, \tag{4.6}$$

where the coefficients z_i at the basis functions of the type (4.3a), (4.3b), (4.3c) or (4.3d) are equal to $z_h(P_i)$ for $P_i \in \sigma^{(1)}$, $z_h(P_i)$ for $P_i \in \sigma^{(2)}$, $z_h \mid \Gamma_{\mathrm{I}}^i$ or $z_h \mid (\Gamma_{\mathrm{T}}^i \cup S_i)$, respectively.

4.2. Algebraic system equivalent to the discrete problem

We seek an approximate solution of the form

$$u_h = u_h^* + z_h, \quad z_h \in \tilde{V}_h, \tag{4.7}$$

that is,

$$u_h = u_h^* + \sum_{j=1}^{n} z_j \tilde{w}_j^*. \tag{4.8}$$

Hence, (3.17) can be written in the equivalent form

$$a_h\left(u_h^* + \sum_{j=1}^{n} z_j \tilde{w}_j^*, w_i^*\right) = L_h(w_i^*), \quad i = 1, \dots, n, \tag{4.9}$$

or, in view of (3.13) and (3.14),

$$\mathbb{A}(\bar{z})\bar{z} = \Phi(\bar{z}), \tag{4.10}$$

where for each $\bar{z} = (z_1, \dots, z_n) \in \mathbb{R}^n$, $\mathbb{A}(\bar{z})$ is an $n \times n$ matrix and $\Phi(\bar{z}) \in \mathbb{R}^n$, $\mathbb{A}(\bar{z}) = (a_{ij}(\bar{z}))_{i,j=1}^n$, $\Phi(\bar{z}) = (\Phi_1(\bar{z}), \dots, \Phi_n(\bar{z}))^{\mathrm{T}}$,

$$a_{ij}(\bar{z}) = \sum_{T \in \mathscr{T}_h} \nabla w_i^* | T \cdot \nabla \tilde{w}_j^* | T I_T(\bar{z}), \tag{4.11a}$$

$$\Phi_i(\bar{z}) = L_h(w_i^*) - \sum_{T \in \mathscr{T}_h} \nabla u_h^* | T \cdot \nabla w_i^* | T I_T(\bar{z}), \tag{4.11b}$$

$$I_T(\bar{z}) = \mathscr{I}_T\left(u_h^* + \sum_{j=1}^{n} z_j \tilde{w}_j^*\right), \tag{4.11c}$$

$$c_1 \leqslant I_T(\bar{z}) \operatorname{meas}^{-1}(T) \leqslant c_2, \quad \bar{z} \in \mathbb{R}^n. \tag{4.11d}$$

4.3. Discrete ~~problem without~~ trailing conditions

Let $K_{\mathrm{T}} = 0$. Then $V_h = \tilde{V}_h$ and the matrix $\mathbb{A}(\bar{z})$ is symmetric for each $\bar{z} \in \mathbb{R}^n$. Moreover, for $\bar{v} = (v_1, \dots, v_n) \neq 0$ and $v_h = \sum_{i=1}^n v_i w_i^*$, we have

$$\sum_{i,j}^{n} a_{ij}(\bar{z}) v_i v_j = \sum_{T \in \mathscr{T}_h} \mathscr{I}_T(u_h^* + z_h) \,|\nabla v_h | T|^2 \geqslant c_1 \int_{\Omega_h} |\nabla v_h|^2 \,\mathrm{d}x > 0, \tag{4.12}$$

as it follows from (1.19), (3.11), (3.14) and the fact that $v_h \not\equiv$ const. in $\bar{\Omega}_h$ ($v_h = 0$ on $\bar{\Gamma}_{\mathrm{D}h}$). It means that $\mathbb{A}(\bar{z})$ is positive definite.

With the use of the ideas from the monotone operator theory it is possible to prove the unique solvability of the discrete problem and the convergence of approximate solutions to the exact solution of problem (2.7) (see [14,16]). By the discretization of (2.9) we get an iterative process for obtaining the approximate solution:

$$\tilde{\mathbb{B}}\bar{z}^0 = \Psi, \tag{4.13a}$$

$$\mathbb{B}\bar{\xi}^{k+1/2} = -(\mathbb{A}(\bar{z}^k)\bar{z}^k - \Phi(\bar{z}^k)), \quad \bar{z}^{k+1} = \bar{z}^k + \nu\bar{\xi}^{k+1/2}, \ k \geqslant 0. \tag{4.13b}$$

Here $\tilde{\mathbb{B}} = \mathbb{B}$ is the symmetric positive definite $n \times n$ *preconditioning matrix* with the elements

$$b_{ij} = \sum_{T \in \mathscr{T}_h} \mathscr{I}_T(0)\, \nabla w_i^* | T \cdot \nabla w_j^* | T, \tag{4.14}$$

and $\Psi \in \mathbb{R}^n$ is the vector with components

$$\Psi_i = L_h(w_i^*) - \sum_{T \in \mathscr{T}_h} \mathscr{I}_T(0)\, \nabla u_h^* | T \cdot \nabla w_i^* | T. \tag{4.15}$$

It means that (4.13a) is a system obtained by the discretization of the incompressible flow problem. In the same way as in [13] we prove the existence of $\tilde{\nu} > 0$ such that for each $\nu \in (0, \tilde{\nu})$ the sequence $\bar{z}^k$ converges to the solution of system (4.10).

4.4. Properties of the general discrete problem

The purpose of the following considerations will be the investigation of the unique solvability of the discrete problem (3.17). We shall proceed similarly as in [11] devoted to the finite-difference solution of flow problems with trailing conditions.

Let $\psi \in X_h$ and $T \in \mathscr{T}_h$ be a triangle with vertices $P = (p_1, p_2)$, $Q = (q_1, q_2)$, $R = (r_1, r_2)$. Then

$$\psi(x) = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2, \quad x = (x_1, x_2) \in T, \tag{4.16}$$

where

$$\begin{gathered}
\alpha_0 = \frac{1}{D}\begin{vmatrix} p_1 & p_2 & \psi(P) \\ q_1 & q_2 & \psi(Q) \\ r_1 & r_2 & \psi(R) \end{vmatrix}, \qquad \alpha_1 = -\frac{1}{D}\begin{vmatrix} p_2 & \psi(P) & 1 \\ q_2 & \psi(Q) & 1 \\ r_2 & \psi(R) & 1 \end{vmatrix}, \\
\alpha_2 = \frac{1}{D}\begin{vmatrix} p_1 & \psi(P) & 1 \\ q_1 & \psi(Q) & 1 \\ r_1 & \psi(R) & 1 \end{vmatrix}, \qquad D = \begin{vmatrix} p_1 & p_2 & 1 \\ q_1 & q_2 & 1 \\ r_1 & r_2 & 1 \end{vmatrix}, \quad |D| = 2\,\mathrm{meas}(T) \neq 0,
\end{gathered} \tag{4.17}$$

and

$$\nabla\psi \,|\, T = (\alpha_1, \alpha_2). \tag{4.18}$$

Let $T \in \mathscr{T}_h$ be a triangle with vertices P, Q, R and $w \in X_h$. We write

$$w \,|\, T = w_P, \qquad \text{if } w(P) = 1,\ w(Q) = w(R) = 0, \tag{4.19a}$$

$$w \,|\, T = w_{P,Q}, \qquad \text{if } w(P) = w(Q) = 1,\ w(R) = 0, \tag{4.19b}$$

$$w \,|\, T = w_{P,Q,R,\alpha}, \quad \text{if } w(P) = w(Q) = w(R) = \alpha. \tag{4.19c}$$

Lemma 4.1. *Let $T \in \mathscr{T}_h$ be a triangle with vertices $P=(p_1, p_2)$, $Q=(q_1, q_2)$, $R=(r_1, r_2)$. Then*

$$\nabla w_P = \frac{1}{D}(q_2 - r_2,\ r_1 - q_1), \tag{4.20a}$$

$$\nabla w_{P,Q} = \frac{1}{D}(q_2 - p_2,\ p_1 - q_1) = -\nabla w_R | T, \tag{4.20b}$$

$$\nabla w_{P,Q,R,\alpha} = 0, \tag{4.20c}$$

$$|\nabla w_P|^2 = \frac{1}{D^2}|Q-R|^2, \tag{4.20d}$$

$$|\nabla w_P|^2 = -\nabla w_P \cdot \nabla w_Q - \nabla w_P \cdot \nabla w_R, \tag{4.20e}$$

$$\nabla w_P \cdot \nabla w_Q = -\frac{1}{D^2}|P-R|\,|Q-R| \cos \alpha(R), \tag{4.20f}$$

where $\alpha(R)$ is the angle in T at the vertex R.

Proof. This is a consequence of (4.16)–(4.19) applied to the functions w_P, w_Q, w_R etc. and the well-known expression of the cosine of the angle between two vectors. □

By supp v we denote the support of a function v, i.e., supp $v = \overline{\{x;\ v(x) \neq 0\}}$. For the following considerations we introduce some assumptions concerning the triangulation:

$$\mathscr{T}_h \text{ is sufficiently fine}, \tag{4.21}$$

$$\mathscr{T}_h \text{ is } \textit{weakly acute}\text{, i.e., all angles of all } T \in \mathscr{T}_h \text{ are } \leqslant \tfrac{1}{2}\pi. \tag{4.22}$$

(The demands on $\mathscr{T}_h$ representing (4.21) will be cleared in the proof of Theorem 4.2.)

Theorem 4.2. (i) *The elements of the matrix $\mathbb{A}(\bar{z})$, the components of the vector $\Psi(\bar{z})$ and their first-order partial derivatives are continuous and bounded in $\mathbb{R}^n$.*

(ii) *For each $\bar{z} \in \mathbb{R}^n$ the matrix $\mathbb{A}(\bar{z})$ is irreducibly diagonally dominant.*

Proof. The first assertion is a consequence of (4.11), (3.14) and the properties of the function b.

In order to prove the second assertion let us show that

$$a_{ii}(\bar{z}) > 0, \qquad i = 1, \ldots, n, \tag{4.23a}$$

$$a_{ij}(\bar{z}) \leqslant 0 \qquad i, j = 1, \ldots, n,\ i \neq j, \tag{4.23b}$$

$$\sum_{j=1}^{n} a_{ij}(\bar{z}) \geqslant 0, \quad i = 1, \ldots, n, \tag{4.23c}$$

$$\text{there exists } i_0 \in \{1, \ldots, n\} \text{ such that } \sum_{j=1}^{n} a_{i_0 j}(\bar{z}) > 0 \tag{4.23d}$$

hold for all $\bar{z} \in \mathbb{R}^n$.

(4.23a) If $w_i^* = \tilde{w}_i^*$, then by (4.11), (3.14) and (1.19)

$$a_{ii}(\bar{z}) = \sum_{T \in \mathscr{T}_h} |\nabla w_i^* | T|^2 I_T(\bar{z}) \geqslant c_1 \int_{\Omega_h} |\nabla w_i^*|^2 \, \mathrm{d}x > 0. \tag{4.24}$$

The case $w_i^* \neq \tilde{w}_i^*$ corresponds to the basis functions of the type (4.3d) and (4.3$\tilde{\text{d}}$) associated with the vertex $P_i = P_j^*$, $j = 1, \ldots, K_{\text{T}}$. Then,

$$\operatorname{supp} w_i^* \subset \operatorname{supp} \tilde{w}_i^*, \quad w_i^* = w_i,$$
$$\nabla w_i^* = \nabla w_i, \qquad \nabla \tilde{w}_i^* = \nabla w_i + \sum_{P_k \in \sigma_h \cap \bar{\Gamma}_{\text{T}}^j} \nabla w_k,$$

and for each $T \in \mathcal{T}_h$,

$$\nabla w_i^* | T \cdot \nabla \tilde{w}_i^* | T = | \nabla w_i | T |^2 + \sum_{P_k \in \sigma_h \cap \bar{\Gamma}_{\text{T}}^j} (\nabla w_i | T \cdot \nabla w_k | T). \tag{4.25}$$

If $T \not\subset \operatorname{supp} w_i^* = \operatorname{supp} w_i$, then $\nabla w_i^* | T \cdot \nabla \tilde{w}_i^* | T = 0$.

Let $T \subset \operatorname{supp} w_i^*$. Then $P_i \in T \cap \sigma_h$ and at most three terms in the right-hand side of (4.25) are nonzero (corresponding to the vertex P_i and the vertices P_k of T such that $P_k \in \sigma_h \cap \bar{\Gamma}_{\text{T}}^j$). In virtue of (4.22) and (4.20f),

$$\nabla w_i | T \cdot \nabla w_k | T \leqslant 0,$$

for the vertices P_i, $P_k \in T \cap \sigma_h$, $i \neq k$. From this, (4.25) and (4.20e) it follows that

$$\nabla w_i^* | T \cdot \nabla \tilde{w}_i^* | T \geqslant 0. \tag{4.26}$$

Since the triangulation $\mathcal{T}_h$ is sufficiently fine, there exists a triangle $\hat{T} \in \mathcal{T}_h$, $\hat{T} \subset \operatorname{supp} w_i^*$, with vertices P_i $(= P_j^*)$, P_l, $P_m \notin \bar{\Gamma}_{\text{T}}^j$ and thus,

$$\nabla w_i^* \big| \hat{T} \cdot \nabla \tilde{w}_i^* \big| \hat{T} = \big| \nabla w_i \hat{T} \big|^2 > 0.$$

From this, (4.26) and (4.11) we conclude that

$$a_{ii}(\bar{z}) \geqslant c_1 \operatorname{meas}(\hat{T}) | \nabla w_i | \hat{T} |^2 > 0. \tag{4.27}$$

(4.23b) For $i \neq j$, by (4.3),

$$w_i^* = \sum_{P_k \in K_i} w_k, \qquad \tilde{w}_j^* = \sum_{P_l \in L_j} w_l,$$
$$K_i,\ L_j \subset \sigma_h, \quad K_i \cap L_j = \emptyset.$$

In view of (4.20f) and (4.22), we have $\nabla w_k | T \cdot \nabla w_l | T \leqslant 0$ (provided $k \neq l$), which immediately yields

$$\nabla w_i^* | T \cdot \nabla \tilde{w}_j^* | T \leqslant 0, \quad T \in \mathcal{T}_h.$$

Hence, by (4.11),

$$a_{ij}(\bar{z}) \leqslant 0, \quad \text{for } i \neq j. \tag{4.28}$$

(4.23c) From (4.11) we have

$$\sum_{j=1}^{n} a_{ij}(\bar{z}) = \sum_{T \in \mathcal{T}_h} I_T(\bar{z}) \sum_{j=1}^{n} \left(\nabla w_i^* | T \cdot \nabla \tilde{w}_j^* | T \right). \tag{4.29}$$

Let $T \in \mathcal{T}_h$ be an arbitrary triangle with vertices P, Q, R. We set

$$\sigma_{i,T} = \sum_{j=1}^{n} \left(\nabla w_i^* | T \cdot \nabla \tilde{w}_j^* | T \right). \tag{4.30}$$

With respect to possible forms of the functions w_i^*, $\tilde{w}_j^*$ we distinguish the following cases.

(α) $w_i^* \mid T = w_{P,Q,R,\alpha}$, $\alpha = 0$ or $\alpha = 1$. Then $\sigma_{i,T} = 0$.

(β) $w_i^* \mid T = w_{P,Q}$. In view of (4.21) we can assume that $\tilde{w}_i^* \mid T = w_i^* \mid T$ (= basis functions corresponding to $\bar{\Gamma}_{\rm I}^j$). For $j \neq i$ the only nonzero term $\nabla w_i^* \mid T \cdot \nabla \tilde{w}_j^* \mid T$ is obtained, provided $R \notin \sigma_h \cap \bar{\Gamma}_{\rm D}$, $P_j = R$ and $\tilde{w}_j^* \mid T = w_R \mid T$. Then, by (4.20b),

$$\sigma_{i,T} = \left(\nabla w_{P,Q} \mid T\right)^2 + \nabla w_{P,Q} \mid T \cdot \nabla w_R \mid T = 0.$$

If $R \in \sigma_h \cap \bar{\Gamma}_{\rm D}$, then

$$\sigma_{i,T} = \left(\nabla w_{P,Q} \mid T\right)^2 > 0.$$

(γ) $w_i^* \mid T = w_P$. Then we have the following possibilities.

($\gamma\alpha$) $\tilde{w}_i^* \mid T = w_i^* \mid T = w_P \mid T$ and $\tilde{w}_j^* \mid T = w_Q \mid T$ or $w_R \mid T$, provided Q, $R \notin \sigma_h \cap \bar{\Gamma}_{\rm D}$. Then by (4.20e),

$$\sigma_{i,T} = \left(\nabla w_P \mid T\right)^2 + \nabla w_P \mid T \cdot \nabla w_Q \mid T + \nabla w_P \mid T \cdot \nabla w_R \mid T = 0.$$

If, e.g., $Q \in \sigma_h \cap \bar{\Gamma}_{\rm D}$, then by (4.20e), (4.20f) and the fact that $\alpha(R) \leqslant \frac{1}{2}\pi$,

$$\sigma_{i,T} = \left(\nabla w_P \mid T\right)^2 + \nabla w_P \mid T \cdot \nabla w_R \mid T = -\nabla w_P \mid T \cdot \nabla w_Q \mid T \geqslant 0.$$

In the case $\alpha(R) < \frac{1}{2}\pi$ we get $\sigma_{i,T} > 0$. Similar results hold if $R \in \sigma_h \cap \bar{\Gamma}_{\rm D}$ or $Q, R \in \sigma_h \cap \bar{\Gamma}_{\rm D}$.

($\gamma\beta$) $\tilde{w}_i^* \mid T = w_i^* \mid T = w_P \mid T$, $\tilde{w}_j^* \mid T = w_{Q,R} \mid T$, $Q,R \notin \sigma_h \cap \bar{\Gamma}_{\rm D}$. Using (4.20b), where we change P, Q, R to Q, R, P, we get

$$\sigma_{i,T} = \left(\nabla w_P \mid T\right)^2 + \nabla w_P \mid T \cdot \nabla w_{Q,R} \mid T = 0.$$

($\gamma\gamma$) $\tilde{w}_i^* \mid T = w_{P,Q} \mid T$, $\tilde{w}_j^* \mid T = w_R \mid T$, provided $R \notin \sigma_h \cap \bar{\Gamma}_D$. In view of (4.20b), we have

$$\sigma_{i,T} = \left(\nabla w_P \mid T\right) \cdot \nabla w_{P,Q} \mid T + \nabla w_P \mid T \cdot \nabla w_R \mid T = 0.$$

If $R \in \sigma_h \cap \bar{\Gamma}_{\rm D}$, then

$$\sigma_{i,T} = \left(\nabla w_P \mid T\right) \cdot \nabla w_{P,Q} \mid T = -\nabla w_P \mid T \cdot \nabla w_R \mid T \geqslant 0.$$

If $\alpha(Q) < \frac{1}{2}\pi$, then $\sigma_{i,T} > 0$.

($\gamma\delta$) $\tilde{w}_i^* \mid T = w_{P,R} \mid T$, $\tilde{w}_j^* \mid T = w_Q \mid T$. This case is analogous to ($\gamma\gamma$).

As $\Gamma_{\rm D} \neq \emptyset$ and $\mathscr{T}_h$ satisfies (4.21), (4.22), we can assume that there exist vertices $P_{i_0} = P \in \Omega_h$, Q, R of a triangle $T \in \mathscr{T}_h$ such that $Q \in \bar{\Gamma}_{\rm D} \cap \sigma_h$ and $\alpha(R) < \frac{1}{2}\pi$. Then $\sigma_{i_0,T} > 0$.

Now, from (4.29), (4.30) and the above results we conclude that (4.23c), (4.23d) hold.

Finally, as $\mathscr{T}_h$ is sufficiently fine, we can assume that the oriented graph of the matrix $\mathbb{A}(\bar{z})$ is strongly connected and thus, $\mathbb{A}(\bar{z})$ is irreducible (cf. [9,24]). This completes the proof. □

Corollary 4.3. *Under the assumptions of Theorem* 4.2 *the matrix* $\mathbb{A}(\bar{z})$ *is regular and monotone for each* $\bar{z} \in \mathbb{R}^n$ (*see* [2,9]).

From the proof of Theorem 4.2 we can derive the following lemma.

Lemma 4.4. *There exist constants* α_{ij}, β_{ij}, ϵ_i *such that*

$$\begin{aligned} &0 < \alpha_{ii} \leqslant \beta_{ii}, \quad \epsilon_i \geqslant 0, && i = 1, \ldots, n, \\ &\alpha_{ij} \leqslant \beta_{ij} \leqslant 0, && i,\ j = 1, \ldots, n,\ i \neq j, \\ &\epsilon_{i_0} > 0, \quad \textit{for some } i_0 \in \{i = 1, \ldots, n\}, && \end{aligned} \tag{4.31}$$

the oriented graph of the matrix

$$\mathscr{A} = \begin{pmatrix} \alpha_{11} & \beta_{12} & \cdots & \beta_{1n} \\ \beta_{21} & \alpha_{22} & \cdots & \beta_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{n1} & \beta_{n2} & \cdots & \alpha_{nn} \end{pmatrix} \tag{4.32}$$

is strongly connected and

$$\begin{aligned} & a_{ij}(\bar{z}) \in [\alpha_{ij}, \beta_{ij}], \quad i, j = 1, \dots, n, \\ & \sum_{j=1}^{n} a_{ij}(\bar{z}) \geqslant \epsilon_i, \quad i = 1, \dots, n, \ \bar{z} \in \mathbb{R}^n. \end{aligned} \tag{4.33}$$

Lemma 4.5. *Let*

$$\mathscr{M} = \Big\{\mathbb{A} = (a_{ij})_{i,j=1}^{n};\ a_{ij} \in [\alpha_{ij}, \beta_{ij}] \text{ for } i, j = 1, \dots, n \text{ and } \sum_{j=1}^{n} a_{ij} \geqslant \epsilon_i \text{ for } i = 1, \dots, n\Big\},$$

where the numbers α_{ij}, β_{ij}, ϵ_i have the properties from Lemma 4.4. *Then there exists $\delta > 0$ such that*

$$|\det \mathbb{A}| \geqslant \delta, \quad \forall \mathbb{A} \in \mathscr{M}. \tag{4.34}$$

Proof. If $\mathbb{A} \in \mathscr{M}$, then the oriented graph of $\mathbb{A}$ contains the oriented graph of $\mathscr{A}$. Hence, the matrix $\mathbb{A}$ is irreducibly diagonally dominant and $|\det \mathbb{A}| > 0$. It is obvious that $\mathscr{M}$ considered as a subset of $\mathbb{R}^{n^2}$ is bounded and closed. Since $m(\mathbb{A}) = |\det \mathbb{A}|$ is a continuous function of the elements a_{ij} of $\mathbb{A}$, m attains its minimum on $\mathscr{M}$, which is necessarily positive. □

Now we come to the fundamental result on the solvability of the discrete problem.

Theorem 4.6. *Let $\mathscr{T}_h$ satisfy* (4.21), (4.22). *Then system* (4.10) *has at least one solution $\bar{z} \in \mathbb{R}^n$, and thus the discrete problem* (3.17) *has at least one solution $u_h \in X_h$.*

Proof. In virtue of Theorem 4.2 and Lemmas 4.4 and 4.5, there exist constants c_1, c_2, $c_3 > 0$ such that

$$\begin{aligned} & \|\mathbb{A}(\bar{z}^{(1)}) - \mathbb{A}(\bar{z}^{(2)})\| \leqslant c_1 \|\bar{z}^{(1)} - \bar{z}^{(2)}\|, \\ & \|\Phi(\bar{z}^{(1)}) - \Phi(\bar{z}^{(2)})\| \leqslant c_1 \|\bar{z}^{(1)} - \bar{z}^{(2)}\|, \quad \bar{z}^{(1)}, \bar{z}^{(2)} \in \mathbb{R}^n, \\ & \|\mathbb{A}^{-1}(\bar{z})\| \leqslant c_2, \quad \|\Phi(\bar{z})\| \leqslant c_3, \qquad \bar{z} \in \mathbb{R}^n. \end{aligned}$$

Here $\|\bar{z}\|$ denotes an arbitrary norm of $\bar{z} \in \mathbb{R}^n$ and $\|\mathbb{A}\|$ is the norm of the matrix $\mathbb{A}$ induced by $\|\bar{z}\|$. This implies that the mapping

$$F(\bar{z}) = \mathbb{A}^{-1}(\bar{z})\Phi(\bar{z})$$

is Lipschitz-continuous and $F(\mathbb{R}^n)$ is a bounded set. By the well-known Brouwer theorem, the mapping F has at least one fixed point $\bar{z} \in \mathbb{R}^n$. Thus, $\bar{z}$ satisfies the relation $\bar{z} = F(\bar{z})$, which is equivalent to (4.10). From this and (4.9) it follows that the function u_h defined by (4.8) is a solution of problem (3.17). □

Further, let us prove the following theorem.

Theorem 4.7. *The solution of problem* (3.17) *is unique.*

Proof. Let u_h^1, u_h^2 be two solutions of (3.17). Then

$$a_h(u_h^2, v_h) - a_h(u_h^1, v_h) = 0, \quad \forall v_h \in V_h.$$

From this, using (3.13), (3.14), the properties of the function b and the mean-value theorem, we get the relation

$$\begin{aligned}
0 = \sum_{T \in \mathscr{T}_k} & \operatorname{meas}(T) \sum_{l=1}^{k_T} \alpha_{T,l} \\
& \times \int_0^1 \Big\{ b\big(x_{T,l}, |\nabla(u_h^1 + t(u_h^2 - u_h^1))|T|^2\big) \\
& \qquad + 2\frac{\partial b}{\partial \eta}\big(x_{T,l}, |\nabla(u_h^1 + t(u_h^2 - u_h^1))|T|^2\big) \big|\nabla(u_h^1 + t(u_h^2 - u_h^1))|T\big|^2 \Big\} \, \mathrm{d}t \\
& \times \nabla(u_h^2 - u_h^1)|T \cdot v_h|T, \quad \forall v_h \in V_h.
\end{aligned} \tag{4.35}$$

Let us set

$$\tilde{I}_T = \sum_{l=1}^{k_T} \alpha_{T,l} \int_0^1 \{\cdots\} \, \mathrm{d}t, \quad T \in \mathscr{T}_h,$$

where $\{\cdots\}$ denotes the expression between braces in the right-hand side of (4.35). In view of (1.19)–(1.21),

$$0 < c_1 \leqslant \tilde{I}_T \leqslant c_1 + 2c_4, \quad T \in \mathscr{T}_h. \tag{4.36}$$

In (4.35) let us set $v_h := w_i^*$, $i = 1, \ldots, n$, and realize that in virtue of (3.17b), $u_h^2 - u_h^1 \in V_h$, which means that

$$u_h^2 - u_h^1 = \sum_{j=1}^{n} \alpha_j \tilde{w}_j^*.$$

Hence (4.35) is equivalent to the system

$$\sum_{j=1}^{n} c_{ij} \alpha_j = 0, \quad i = 1, \ldots, n, \tag{4.37}$$

where

$$c_{ij} = \sum_{T \in \mathscr{T}_h} \operatorname{meas}(T)\, \tilde{I}_T\, \nabla w_i^* | T \cdot \nabla \tilde{w}_j^* | T.$$

In the same way as in Theorem 4.2, using (4.36), we prove that the matrix $(c_{ij})_{i,j=1}^{n}$ is irreducibly diagonally dominant, and hence, regular. Therefore, $\alpha_j = 0$ for $j = 1,\dots,n$, which already implies that $u_h^1 = u_h^2$. $\square$

4.5. Approximate solution

The solution of the problem with trailing conditions is calculated with the use of the iterative process (4.13), where the matrix $\mathbb{B}$ is defined by (4.14) and $\tilde{\mathbb{B}}$ is the matrix with the elements

$$\tilde{b}_{ij} = \sum_{T\in\mathscr{T}_h} \mathscr{I}_T(0)\, \nabla w_i^* | T \cdot \nabla \tilde{w}_j^* | T. \tag{4.38}$$

Under the above assumptions $\tilde{\mathbb{B}}$ is (nonsymmetric) irreducibly diagonally dominant.

Series of numerical experiments proved the convergence of the process (4.13) also in this nonsymmetric case.

5. Convergence

In order to support our numerical results by the investigation of the convergence of the approximate finite-element solutions to the exact one, we shall confine our considerations to a linear model problem of a plane incompressible flow past an isolated profile mentioned in Section 1.1, where we shall assume that $\Gamma_\mathrm{N} = \emptyset$ and $b \equiv 1$. Here we shall introduce the sketch of the L^∞-convergence proof based on Gershgorin's method often used in the connection with finite-difference techniques [2,9,11].

5.1. Continuous model problem

Let Ω be a bounded domain with a sufficiently smooth boundary formed by two components: Γ_T^1, a smooth profile (= inner part of $\partial\Omega$) and Γ_D (= outer part of the boundary). We seek a function $u : \overline{\Omega} \to \mathbb{R}^1$ and a constant $q_\mathrm{T}^1 \in \mathbb{R}^1$ such that

$$-\Delta u = 0, \quad \text{in } \Omega, \tag{5.1a}$$

$$u \mid \Gamma_\mathrm{D} = u_\mathrm{D}, \tag{5.1b}$$

$$u \mid \Gamma_\mathrm{T}^1 = q_\mathrm{T}^1, \tag{5.1c}$$

$$\frac{\partial u}{\partial n}(z_\mathrm{T}^1) = 0. \tag{5.1d}$$

Here u_D is a given function, $z_\mathrm{T}^1 \in \Gamma_\mathrm{T}^1$ is a given point.

We also consider the following problem.

5.2. Auxiliary problem

Find $w : \overline{\Omega} \to \mathbb{R}^1$ such that

$$-\Delta w = f, \quad \text{in } \Omega, \tag{5.2a}$$

$$w \mid \Gamma_\mathrm{D} = w_\mathrm{D}, \tag{5.2b}$$

$$w \mid \Gamma_{\mathrm{T}}^{1} = \text{const.}, \tag{5.2c}$$

$$\frac{\partial w}{\partial n}(z_{\mathrm{T}}^{1}) = v_{\mathrm{T}}^{j}, \tag{5.2d}$$

where $f : \bar{\Omega} \to \mathbb{R}^1$ and $w_{\mathrm{D}} : \Gamma_{\mathrm{D}} \to \mathbb{R}^1$ are given continuous functions and $v_{\mathrm{T}}^{j} \in \mathbb{R}^1$ is a given constant. We see that (5.1) is a special case of (5.2).

5.3. Triangulations

Let us consider a system $\{\mathscr{T}_h\}_{h \in (0,h_0)}$, $h_0 > 0$, of triangulations of polygonal approximations Ω_h of Ω with properties from Section 3. Let the system $\{\mathscr{T}_h\}_{h \in (0,h_0)}$ be *regular* in the following sense.

$$\theta_h \geqslant \theta_0 > 0, \quad \forall h \in (0, h_0); \tag{5.3a}$$

all sides of all $T \in \mathscr{T}_h$, $h \in (0, h_0)$, are parallel either to the axis x_1 or to the axis x_2 or to the line $x_2 = x_1$, except at most triangles $T \in \mathscr{T}_h$ which have at least one vertex lying on $\partial\Omega$; (5.3b)

conditions (4.21) and (4.22) are satisfied for each $\mathscr{T}_h$, $h \in (0, h_0)$; (5.3c)

$\mathscr{T}_h$ is quasiuniform, i.e., each $T \in \mathscr{T}_h$ contains a circle of radius $c_1 h$ and is contained in a circle of radius $c_2 h$, where the constants $c_1 > 0$ and $c_2 < \infty$ are independent of T and h; (5.3d)

there exists a constant $\epsilon_0 \in (0, \frac{1}{2}\pi)$ independent of h such that the following holds: to each vertex $P_i \in \sigma_h \cap \Omega_h$ of a triangle $\hat{T} \in \mathscr{T}_h$ with a vertex $Q \in \partial\Omega$ and at least one side S parallel neither to x_1 nor to x_2 nor to the line $x_2 = x_1$, there exists $T \in \mathscr{T}_h$ with vertices P_i, P_j, P_k such that $P_j \in \partial\Omega$ and the angle $\alpha(P_k)$ of T at the vertex P_k satisfies the condition $\alpha(P_k) \leqslant \frac{1}{2}\pi - \epsilon_0$. (5.3e)

With respect to (5.3b) we write $\sigma_h \cap \Omega_h = \sigma_h^1 \cup \sigma_h^2$, where $P_i \in \sigma_h^2$ if and only if there exists $T \in \mathscr{T}_h$ with vertices P_i, P_j, P_k, such that $P_j \in \partial\Omega$ and $\sigma_h^1 = \sigma_h \cap \Omega_h \setminus \sigma_h^2$. That is, σ_h^1 is formed by all vertices of all $T \in \mathscr{T}_h$ such that $T \cap \partial\Omega = \emptyset$. The vertices of σ_h^1 are called *regular*. The vertices from σ_h^2 satisfy (5.3e). From (5.3b) it follows that the support $A_i = \operatorname{supp} w_i$ of the basis function $w_i \in X_h$ associated with $P_i \in \sigma_h^1$ has the form drawn in Fig. 5.1.

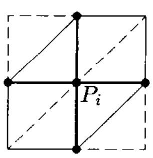

Fig. 5.1.

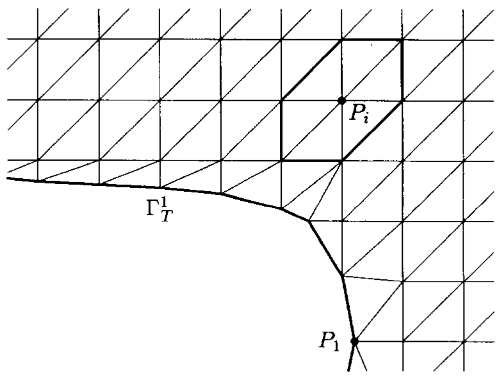

Fig. 5.2.

Similarly as in Section 3, by the discretization of problem (5.2), we get the following problem.

5.4. Auxiliary discrete problem

Let us define $w_h^* \in X_h$, such that

$$\begin{aligned} &w_h^*(P_i) = w_D(P_i), \quad P_i \in \sigma_h \cap \Gamma_D, \\ &w_h^* \mid \Gamma_{Th}^1 = \left|\tilde{P}_1 - P_1^*\right| v_T^1, \\ &w_h^*(P_i) = 0, \quad P_i \in \sigma_h \cap \Omega_h \end{aligned} \tag{5.4}$$

(cf. (3.6)).

Further we consider the form a_h and the spaces V_h and $\tilde{V}_h$ given by (3.13), (3.2) and (3.7), respectively (where we take into account that $K_T = 1$, $K_I = 0$, $\Gamma_N = \Gamma_P^- = \Gamma_P^+ = \emptyset$, $\mathscr{I}_T = \text{meas}(T)$).

Now we seek $w_h : \overline{\Omega}_h \to \mathbb{R}^1$ such that

$$w_h \in X_h, \tag{5.5a}$$

$$w_h - w_h^* \in \tilde{V}_h, \tag{5.5b}$$

$$a_h(w_h, v_h) = \int_{\Omega_h} f v_h \, dx, \quad \forall v_h \in V_h. \tag{5.5c}$$

We see that w_h is given by its values $w_h(P_i)$ for $P_i \in \sigma_h \cap \Omega_h$.

Let the vertices be numbered in such a way that $\sigma_h \cap \Omega_h = \{P_1, \ldots, P_n\}$. (Of course, $n = n_h \to \infty$, if $h \to 0+$.) Then (5.5) is equivalent to the system

$$a_h(w_h, w_i^*) = \int_{\Omega_h} f w_i^* \, dx, \quad i = 1, \ldots, n, \tag{5.6}$$

which can be written in the form

$$\mathbb{A}_h \overline{w}_h = \theta_h. \tag{5.7}$$

Here $\mathbb{A}_h$ is an $n \times n$ matrix, $\mathbb{A}_h = (a_{ij}^h)_{i,j=1}^n$,

$$a_{ij}^h = \sum_{T \in \mathscr{T}_h} \nabla w_i^* | T \cdot \nabla \tilde{w}_j^* | T \ \text{meas}(T), \tag{5.8}$$

$\theta_h = (\theta_{h1}, \ldots, \theta_{hn}) \in \mathbb{R}^n$,

$$\theta_{hi} = \int_{\Omega_h} f w_i^* \, dx - \sum_{T \in \mathscr{T}_h} \nabla w_h^* | T \cdot \nabla w_i^* | T \ \text{meas}(T) \tag{5.9}$$

and $\overline{w}_h = (w_h(P_1), \ldots, w_h(P_n))$. In virtue of Corollary 4.3, the *matrix* $\mathbb{A}_h$ *is monotone*.

5.5. Discrete problem to (5.1)

Find $u_h : \overline{\Omega}_h \to \mathbb{R}^1$ such that

$$u_h \in X_h, \tag{5.10a}$$

$$u_h - u_h^* \in \tilde{V}_h, \tag{5.10b}$$

$$a_h(u_h, v_h) = 0, \quad \forall v_h \in V_h, \tag{5.10c}$$

where $u_h^* \in X_h$, $u_h^*(P_i) = u_D(P_i)$ for $P_i \in \sigma_h \cap \Gamma_D$ and $u_h^*(P_i) = 0$ for $P_i \in \sigma_h \backslash \Gamma_D$. Equation (5.10) is equivalent to the system

$$\mathbb{A}_h \bar{u}_h = \Phi_h, \tag{5.11}$$

where $\bar{u}_h = (u_h(P_1), \ldots, u_h(P_n))$ and $\Phi_h = (\Phi_{h1}, \ldots, \Phi_{hn})$,

$$\Phi_{hi} = - \sum_{T \in \mathcal{T}_h} \nabla u_h^* | T \cdot \nabla w_i^* | T \text{ meas}(T). \tag{5.12}$$

(We notice that $w_i^* = w_i$ for all $i = 1, \ldots, n$.)

5.6. Discrete error of the method

This is defined as the vector $\bar{\eta}_h = (\eta_{h1}, \ldots, \eta_{hn})$, with the components

$$\bar{\eta}_{hi} = u(P_i) - u_h(P_i), \quad i = 1, \ldots, n. \tag{5.13}$$

In the sequel we shall prove the error estimate in *the discrete L^∞-norm*, i.e.,

$$\| \bar{\eta}_h \|_{L_h^\infty} = \max_{P_i \in \sigma_h \cap \Omega_h} | \eta_{hi} | = \max_{\sigma_h \cap \Omega_h} | u - u_h |. \tag{5.14}$$

The process of estimating $\| \bar{\eta}_h \|_{L_h^\infty}$ is the following.

(1) We set $f \equiv 1$, $w_D \equiv 1$, $v_T^1 = 1$ and substitute the values of the exact solutions $w(P_i)$ and $u(P_i)$, $P_i \in \sigma_h \cap \Omega_h$, into the systems (5.7) and (5.11), respectively. With the notation $\bar{\bar{w}}_h = (w(P_1), \ldots, w(P_n))$, $\bar{\bar{u}}_h = (u(P_1), \ldots, u(P_n))$, we get the relations

$$\mathbb{A}_h \bar{\bar{u}}_h = \Phi_h + \epsilon_h, \tag{5.15}$$

$$\mathbb{A}_h \bar{\bar{w}}_h = \theta_h + \tau_h. \tag{5.16}$$

The quantities $\epsilon_h = (\epsilon_{h1}, \ldots, \epsilon_{hn})$ and $\tau_h = (\tau_{h1}, \ldots, \tau_{hn})$ are called the *discretization errors*.

(2) As $\bar{\eta}_h = \bar{\bar{u}}_h - \bar{u}_h$, in view of (5.11) and (5.15),

$$\mathbb{A}_h \bar{\eta}_h = \epsilon_h. \tag{5.17}$$

(3) Provided w, u are smooth enough, we prove the existence of constants h_1, K, $\delta > 0$, $h_1 \leqslant h_0$ (independent of h) such that

$$| \epsilon_{hi} | \leqslant K h^\delta (\theta_{hi} + \tau_{hi}), \quad \forall P_i \in \sigma_h \cap \Omega_h, \ \forall h \in (0, h_1). \tag{5.18}$$

(4) From (5.16) we get

$$\mathbb{A}_h \left(K h^\delta \bar{\bar{w}}_h \right) = K h^\delta (\theta_h + \tau_h).$$

This, (5.17), (5.18), the monotony of $\mathbb{A}_h$ and [2, Theorem 4.11] yield the inequality

$$| \eta_{hi} | \leqslant K h^\delta w(P_i) \leqslant K h^\delta \max_{\bar{\Omega}} | w | = M h^\delta, \quad P_i \in \sigma_h \cap \Omega_h, \ h \in (0, h_1).$$

Hence,

$$\| \bar{\eta}_h \|_{L_h^\infty} \leqslant M h^\delta, \quad h \in (0, h_1), \tag{5.19}$$

which is the sought estimate.

In order to prove (5.18), we must estimate the discretization errors τ_h and ϵ_h and derive the estimate of θ_{hi} from below. We shall use several auxiliary concepts and assertions. For simplicity we denote by c, C generic positive constants independent of h and P_i with different values at different places in general.

Lemma 5.1. *There exist constants* c, $C > 0$ *such that*
(a) $ch^2 \leqslant \text{meas}(T) \leqslant Ch^2$, $\forall T \in \mathcal{T}_h$, $\forall h \in (0, h_0)$;
(b) $ch^2 \leqslant \text{meas}(A_i) \leqslant Ch^2$, $\forall P_i \in \sigma_h$, $\forall h \in (0, h_0)$;
(c) $ch^{-1} \leqslant |\nabla w_i | T| \leqslant Ch^{-1}$, $\forall T \subset A_i$, $\forall P_i \in \sigma_h$, $\forall h \in (0, h_0)$;
(d) $ch^2 \leqslant \int_{\Omega_h} w_i \,\mathrm{d}x \leqslant Ch^2$, $\forall P_i \in \sigma_h$, $\forall h \in (0, h_0)$.

Proof. This is a consequence of (5.3a), (5.3d). □

Let us define the operator of the Lagrange interpolation r_h:

$$r_h v \in X_h, \qquad (r_h v)(P_j) = v(P_j), \quad \forall P_j \in \sigma_h, \tag{5.20}$$

defined for all functions $v : \sigma_h \to \mathbb{R}^1$.

Further, we introduce the interpolation operator $R_h^{v_T^1}$ associated with the trailing condition (5.2d):

$$\begin{aligned} &R_h^{v_T^1} v \in X_h, \qquad \left(R_h^{v_T^1} v\right)(P_i) = r_h v(P_i), \quad \forall P_i \in \sigma_h \cap (\Omega_h \cup \Gamma_D), \\ &R_h^{v_T^1} v(P_i) = v(P_1^*) + \left|\tilde{P}_1 - P_1^*\right| v_T^1, \quad \forall P_i \in \sigma_h \cap \Gamma_T^1, \end{aligned} \tag{5.21}$$

where $v : \sigma_h \to \mathbb{R}^1$.

It is evident that for the solution w of problem (5.2), we have

$$R_h^{v_T^1} w - w_h^* \in \tilde{V}_h. \tag{5.22}$$

Further, (5.16) is equivalent to the system

$$a_h\left(R_h^{v_T^1} w, w_i^*\right) = \int_{\Omega_h} f w_i^* \,\mathrm{d}x + \tau_{hi}, \quad i = 1, \ldots, n. \tag{5.23}$$

In the following we shall set $f \equiv 1$, $w_D \equiv 1$ and $v_T^1 = 1$ and assume that

$$w, u \in C^3(\tilde{\Omega}), \tag{5.24}$$

where $\tilde{\Omega}$ is an open set such that $\overline{\Omega} \cap \overline{\Omega}_h \subset \tilde{\Omega}$ for all $h \in (0, h_0)$. Let us deal with the study of the character of θ_h and τ_h.

(a) If $P_i \in \sigma_h^1$, then (5.6) represents the regular five-point scheme, as it follows from (5.8), (4.20f) and the fact that $w_i^* = \tilde{w}_i^* = w_i$, see Fig. 5.2.

By the use of the Taylor formula we easily prove that the substitution of $w \in C^3(\tilde{\Omega})$ into (5.23) leads to the discretization error τ_{hi} that satisfies the estimate

$$|\tau_{hi}| \leqslant ch^3, \quad \forall P_i \in \sigma_h^1, \ \forall h \in (0, h_0). \tag{5.25}$$

Moreover, Lemma 5.1(d) implies that

$$\theta_{hi} \geqslant Ch^2, \quad \forall P_i \in \sigma_h^1, \ \forall h \in (0, h_0). \tag{5.26}$$

(The constants c, C are independent of h and P_i.)

For the following analysis we put $\tilde{f} = -\Delta w$ in $\tilde{\Omega}$. In view of (5.2a) and (5.24), we have $\tilde{f} \in C^1(\tilde{\Omega})$ and $\tilde{f} = 1$ in $\bar{\Omega}$. By Green's theorem,

$$\int_{\Omega_h} \nabla w \cdot \nabla w_i^* \, \mathrm{d}x = \int_{\Omega_h} \tilde{f} w_i^* \, \mathrm{d}x = \int_{A_i} w_i^* \, \mathrm{d}x - \int_{A_i \setminus \Omega} (\tilde{f} - 1) w_i^* \, \mathrm{d}x. \tag{5.27}$$

Using analogous techniques as in [16], we can prove the existence of a constant $c > 0$ such that

$$\left| \int_{A_i \setminus \Omega} (\tilde{f} - 1) w_i^* \, \mathrm{d}x \right| \leqslant ch^5, \quad \forall P_i \in \sigma_h, \ \forall h \in (0, h_0). \tag{5.28}$$

(b) Let $P_i \in \sigma_h^2$ lie in the neighbourhood of Γ_{D}. It means that P_i satisfies (5.3e) with $P_j \in \Gamma_{\mathrm{D}}$. In virtue of (4.21), $R_h^1 w \mid A_i = r_h w \mid A_i$ and thus,

$$a_h(R_h^1 w, w_i^*) = \int_{A_i} \nabla r_h w \cdot \nabla w_i^* \, \mathrm{d}x. \tag{5.29}$$

Using (5.23) (where $f = 1$, $v_{\mathrm{T}}^1 = 1$), (5.27) and (5.29), we find out that

$$\tau_{hi} = \int_{A_i} (\nabla r_h w - \nabla w) \cdot \nabla w_i^* \, \mathrm{d}x + \int_{A_i \setminus \Omega} (\tilde{f} - 1) w_i^* \, \mathrm{d}x. \tag{5.30}$$

We have

$$\left| \int_{A_i} (\nabla r_h w - \nabla w) \cdot \nabla w_i^* \, \mathrm{d}x \right| \leqslant \| \nabla r_h w - \nabla w \|_{L^\infty(\Omega_h)} \| \nabla w_i^* \|_{L^1(A_i)}. \tag{5.31}$$

Now, well-known approximation results [4, Section 3.1] yield the estimate

$$\| \nabla r_h w - \nabla w \|_{L^\infty(\Omega_h)} \leqslant ch, \quad h \in (0, h_0). \tag{5.32}$$

Further, by Lemma 5.1,

$$\| \nabla w_i^* \|_{L^1(A_i)} \leqslant ch, \quad h \in (0, h_0). \tag{5.33}$$

From (5.28), (5.30)–(5.33) we get the estimate

$$| \tau_{hi} | \leqslant ch^2, \quad h \in (0, h_0), \tag{5.34}$$

(with a constant c independent of P_i and h).

Now let us estimate θ_{hi} from below. By (5.4) and (5.9) (where $f = 1$, $w_{\mathrm{D}} = 1$), we have

$$\theta_{hi} = \int_{A_i} w_i^* \, \mathrm{d}x - \sum_{T \subset A_i} \sum_{P_j \in \sigma_h \cap \Gamma_{\mathrm{D}} \cap A_i} \operatorname{meas}(T) \, \nabla w_i \mid T \cdot \nabla w_j \mid T. \tag{5.35}$$

If we use (4.20f), Lemma 5.1(a) and (5.3e) and take into account Lemma 5.1(d), we prove the existence of a constant $C > 0$ independent of P_i and h such that

$$\theta_{hi} \geqslant C, \quad h \in (0, h_0). \tag{5.36}$$

(c) Let us consider $P_i \in \sigma_h^2$ lying near Γ_{T}^1. Thus, P_i satisfies (5.3e) with $P_j \in \Gamma_{\mathrm{T}}^1$. In view of (5.23), where $f = 1$, $v_{\mathrm{T}}^1 = 1$, and (5.27),

$$\tau_{hi} = \int_{\Omega_h} (\nabla R_h^1 w - \nabla w) \cdot \nabla w_i^* \, \mathrm{d}x + \int_{A_i \setminus \Omega} (\tilde{f} - 1) w_i^* \, \mathrm{d}x. \tag{5.37}$$

Hence,

$$|\tau_{hi}| \leqslant \mathrm{I} + \mathrm{II} + \mathrm{III}, \tag{5.38}$$

$$\mathrm{I} = \left| \int_{A_i} (\nabla r_h w - \nabla w) \cdot \nabla w_i^* \, \mathrm{d}x \right|, \qquad \mathrm{II} = \left| \int_{A_i} (\nabla R_h^1 w - \nabla r_h w) \cdot \nabla w_i^* \, \mathrm{d}x \right|,$$

$$\mathrm{III} = \left| \int_{A_i \setminus \Omega} (\tilde{f} - 1) w_i^* \, \mathrm{d}x \right|.$$

The term I is estimated in the same way as above:

$$0 \leqslant \mathrm{I} \leqslant ch^2, \quad h \in (0, h_0), \tag{5.39}$$

the expression III satisfies (5.28).

Let us deal with II. In virtue of (5.20) and (5.21),

$$R_h^1 w - r_h w = - \sum_{P_j \in \sigma_h \cap \Gamma_\mathrm{T}^1} w_j \left[(w \mid \Gamma_\mathrm{T}^1 - w(P_1^*)) - |\tilde{P}_1 - P_1^*| \right]. \tag{5.40}$$

By (3.4) and (5.2d), there exists a constant $c > 0$ such that

$$\left| (w \mid \Gamma_\mathrm{T}^1 - w(P_1^*)) - |\tilde{P}_1 - P_1^*| \right| \leqslant ch^2, \quad h \in (0, h_0). \tag{5.41}$$

Then, using Lemma 5.1(b, c), we get

$$\| \nabla (R_h^1 w - r_h w) \|_{L^\infty(A_i)} \leqslant ch, \quad h \in (0, h_0). \tag{5.42}$$

Now, from (5.33), the relation

$$0 \leqslant \mathrm{II} \leqslant \| \nabla (R_h^1 w - r_h w) \|_{L^\infty(A_i)} \| \nabla w_i^* \|_{L^1(A_i)}$$

and (5.42) we derive the estimate

$$0 \leqslant \mathrm{II} \leqslant ch^2, \quad h \in (0, h_0). \tag{5.43}$$

Finally, by (5.28), (5.39), (5.43) and (5.38), we have

$$|\tau_{hi}| \leqslant ch^2, \quad h \in (0, h_0), \tag{5.44}$$

with c independent of P_i and h.

In order to estimate θ_{hi} from below, we write (see (5.9))

$$\theta_{hi} = \theta_{hi}^1 + \theta_{hi}^2,$$

$$\theta_{hi}^1 = \int_{\Omega_h} w_i^* dx, \qquad \theta_{hi}^2 = - \sum_{T \in \mathscr{T}_h} \nabla w_h^* | T \cdot \nabla w_i^* | T \ \mathrm{meas}(T). \tag{5.45}$$

Taking into account that $w_h^* \mid \Gamma_\mathrm{T}^1 = |\tilde{P}_1 - P_1^*|$ and $w_h^*(P_k) = 0$ for $P_k \in \sigma_h \cap (A_i \setminus \Gamma_\mathrm{T}^1)$ (cf. (4.21)), we can write

$$w_h^* \mid A_i = \sum_{P_j \in \sigma_h \cap \Gamma_\mathrm{T}^1 \cap A_i} |\tilde{P}_1 - P_1^*| w_j$$

and

$$\theta_{hi}^2 = \sum_{T \in \mathscr{T}_h} \sum_{P_j \in \sigma_h \cap \Gamma_\mathrm{T}^1 \cap A_i} \left[-|\tilde{P}_1 - P_1^*| \ \nabla w_j \mid T \cdot \nabla w_i \mid T \ \mathrm{meas}(T) \right]. \tag{5.46}$$

In virtue of (4.22) and (4.20f) all terms in square brackets in the right-hand side of (5.46) are nonnegative. From (5.3d), (5.3e) and (4.20f), it follows that there exist $\hat{T} \subset A_i$ and $P_j \in \sigma_h \cap \Gamma_T^1 \cap \hat{T}$ such that

$$-|\tilde{P}_1 - P_1^*| \nabla w_j|\hat{T} \cdot \nabla w_i|\hat{T} \text{ meas}(\hat{T}) \geqslant Ch.$$

Hence, $\theta_{hi}^2 \geqslant Ch$. If we take into account Lemma 5.1(d), we immediately get the existence of $C > 0$, independent of P_i and h, such that

$$\theta_{hi} \geqslant Ch, \quad h \in (0, h_0). \tag{5.47}$$

Now we draw our attention to the estimates of the discretization errors ϵ_{hi}. Since we assume that $u \in C^3(\tilde{\Omega})$ and problem (5.1) is a special case of (5.2), ϵ_{hi} behaves analogously as τ_{hi}. Therefore, there exists a constant $c > 0$, independent of P_i and h, such that (see (5.25), (5.34), (5.44))

$$|\epsilon_{hi}| \leqslant ch^3, \quad \text{for } P_i \in \sigma_h^1, \tag{5.48a}$$

$$|\epsilon_{hi}| \leqslant ch^2, \quad \text{for } P_i \in \sigma_h^2, \tag{5.48b}$$

$h \in (0, h_0)$, $1 \leqslant q < 2$.

Comparing the following inequalities:
(5.48a) with (5.25), (5.26) for $P_i \in \sigma_h^1$,
(5.48b) with (5.34), (5.36) for $P_i \in \sigma_h^2$ lying near Γ_D,
(5.48b) with (5.44), (5.47) for $P_i \in \sigma_h^2$ lying near Γ_T^1,
we easily find out that there exist constants K, h_1, such that (5.18) holds with $\delta = 1$. Thus, the sought estimate (5.19) is valid. Our results can be summarized as the following theorem.

Theorem 5.2. *Let* (5.3) *and* (5.24) *be satisfied. Let u_h be the solution of the discrete problem* (5.10). *Then there exist constants M, $h_1 > 0$ such that the discrete L^∞-norm of the error $\overline{\eta}_h = (u - u_h)|(\sigma_h \cap \Omega_h)$ satisfies the estimate*

$$\|\overline{\eta}_h\|_{L_h^\infty} = \max_{\sigma_h \cap \Omega_h} |u - u_h| \leqslant Mh, \quad h \in (0, h_1). \tag{5.49}$$

Remark 5.3. L^∞-estimates based on the maximum principle were studied in [5]. However, the approach of [5] is not applicable to our problem.

6. Examples

The FORTRAN 77 program package was written for the computer solution of a general problem formulated in Section 1.3. The main parts of this package are the following:
- determination of the domain,
- automatic triangulation of the domain taking into account specific demands following from the unified conception,
- determination of the boundary conditions,
- construction of the discrete problem,
- iterative process for the solution of the discrete problem,
- calculation of the velocity, pressure, density, Mach number etc.,

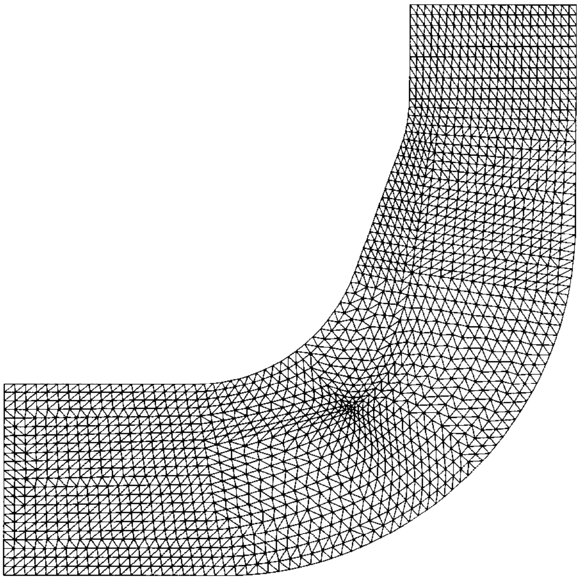

Fig. 6.1.

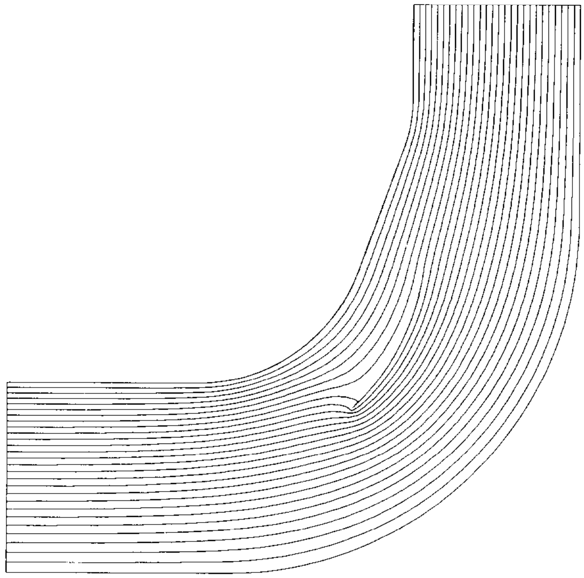

Fig. 6.2.

- determination of isolines of the calculated quantities and their graphs on given curves,
- graphical output.

Linear systems in (4.13) are solved by the well-known SOR-method or the method of conjugate gradients with preconditioning by the symmetric SOR-method (CGSSOR, cf., e.g., [1]).

6.1. Incompressible axially symmetric flow in a channel with an inserted axially symmetric ring

The flow was calculated in a two-dimensional meridional cut of the channel in which the inserted ring was represented by a curve. On the ring the incomplete Dirichlet condition combined with the trailing one was composed.

In Fig. 6.1 we present a triangulation used which consists of 3531 elements. In Figs. 6.2 and 6.3 the isolines of the stream function and isolines of the velocity, respectively, are plotted. In [11] similar results were obtained by the finite-difference method.

6.2. Plane flow past an airfoil

The plane flow past the NACA 0012 airfoil was calculated, where Mach number at infinity $M_\infty = 0.45$, angle of attack $\alpha_\infty = 5°$. The incomplete Dirichlet condition combined with the trailing condition was imposed on the airfoil. The iterative process was stopped after 52 steepest descent iterations (4.13) (10 inner SOR iterations were used again), when the resulting error was 10^{-5}. The relaxation parameter was $\nu = 0.25$.

In Fig. 6.4 we present a triangulation of the domain in which the airfoil was inserted. The triangulation consists of 4234 elements and Fig. 6.5 shows its detail. In Figs. 6.6 and 6.7 the

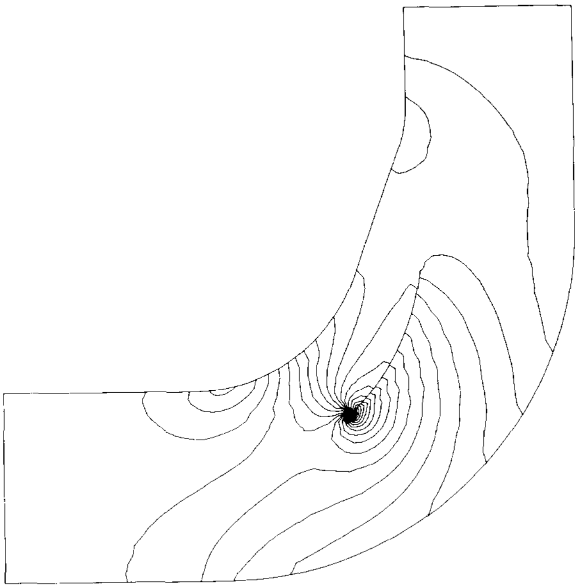

Fig. 6.3.

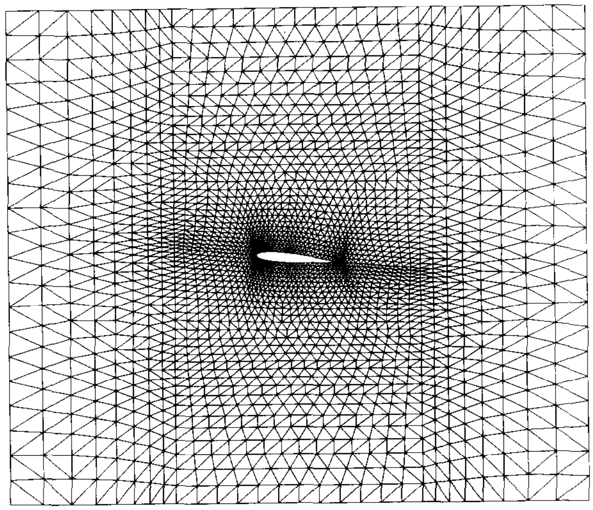

Fig. 6.4.

isolines of the stream function and the isolines of the velocity are plotted. The results are comparable with the results from [6], where another approach was applied.

6.3. Cascade flow

In this example we present the results of the computation of flow past a cascade of profiles $\{C_k\}_{k=-\infty}^{\infty}$ from Fig. 1.3. Figure 6.8 represents the triangulation which consist of 673 vertices and

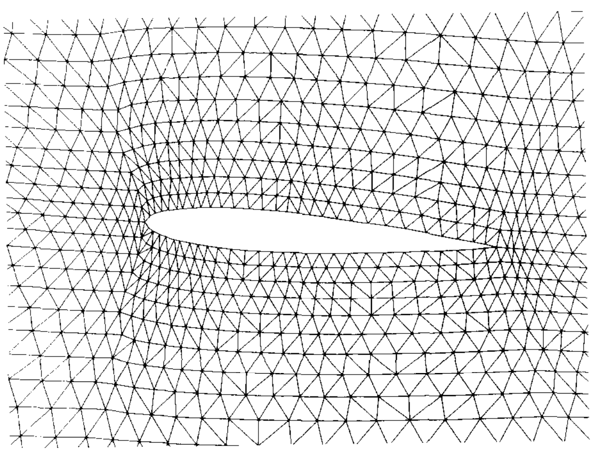

Fig. 6.5.

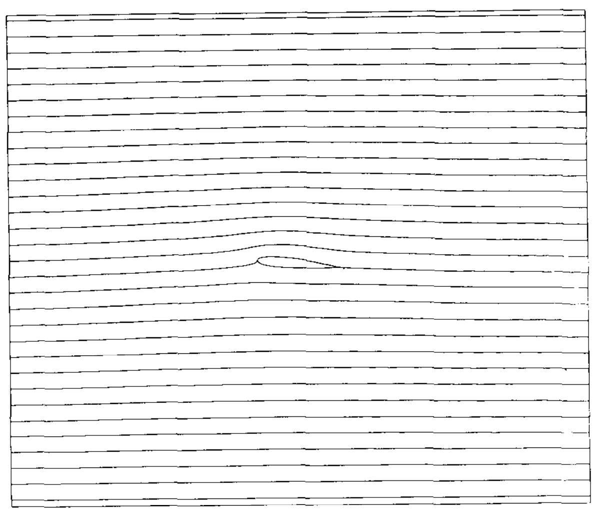

Fig. 6.6.

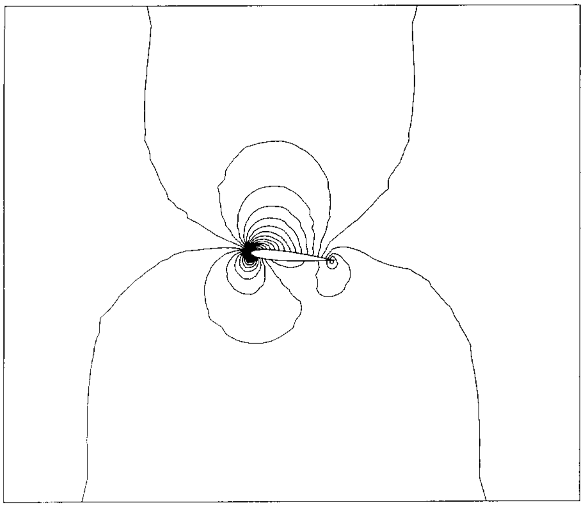

Fig. 6.7.

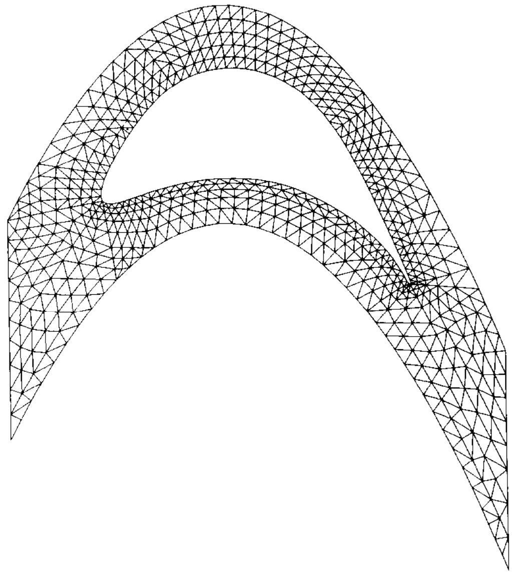

Fig. 6.8.

1137 triangles. In Fig. 6.9 a detail of the triangulation near the trailing point is plotted.

We calculated both incompressible and compressible flow through the cascade for the inlet angle 57° and the inlet velocity $V_1 = 150\ \mathrm{ms}^{-1}$. In the compressible case the Poisson adiabatic constant $\kappa = 1.4$. The density ρ_0 and the speed of sound a_0 at zero velocity have the values $\rho_0 = 1.3\ \mathrm{kgm}^{-3}$ and $a_0 = 330\ \mathrm{ms}^{-1}$, respectively. The corresponding inlet Mach number $M_1 = 0.464$. In the compressible case we tested the influence of the choice of the relaxation parameter ν on the speed of the convergence of the steepest descent iterative process (4.13) (with 10 inner SOR iterations). In Table 6.1 we introduce the number of iterations necessary

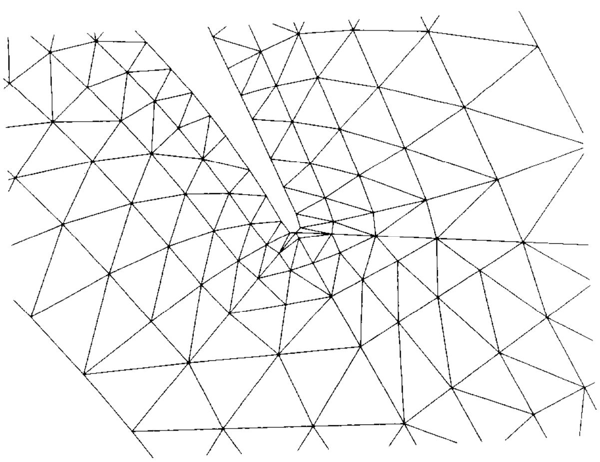

Fig. 6.9.

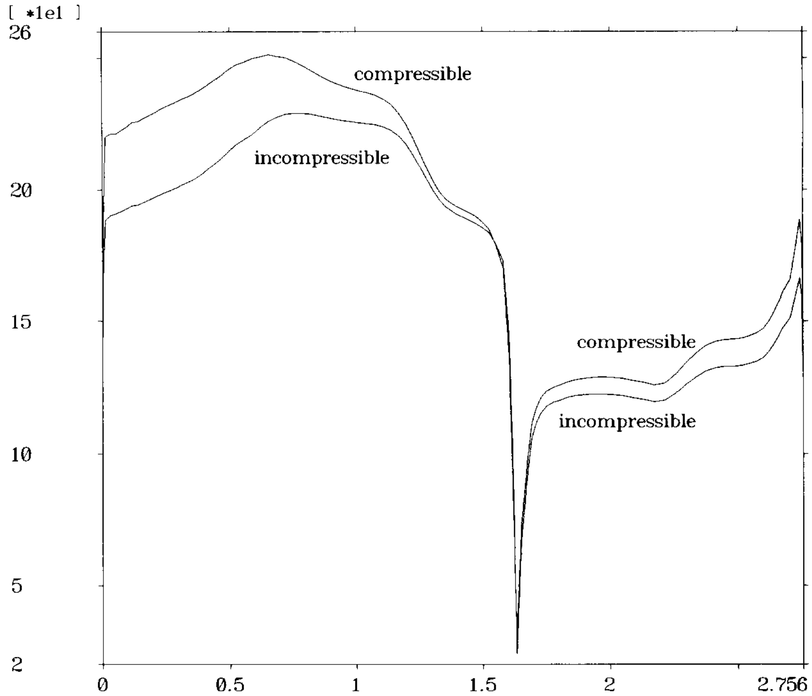

Fig. 6.10. Velocity distribution on the profile.

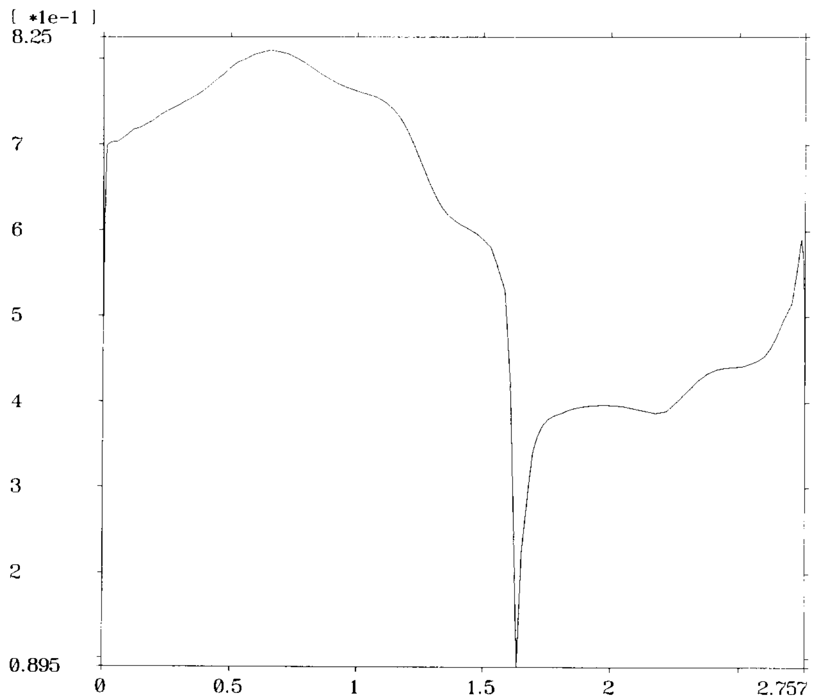

Fig. 6.11. Mach number distribution on the profile.

Table 6.1

Parameter ν	0.25	0.4	0.5	0.6	0.65	0.75	0.775	0.8	0.9
Number of iterations	194	121	97	81	75	65	63	82	82

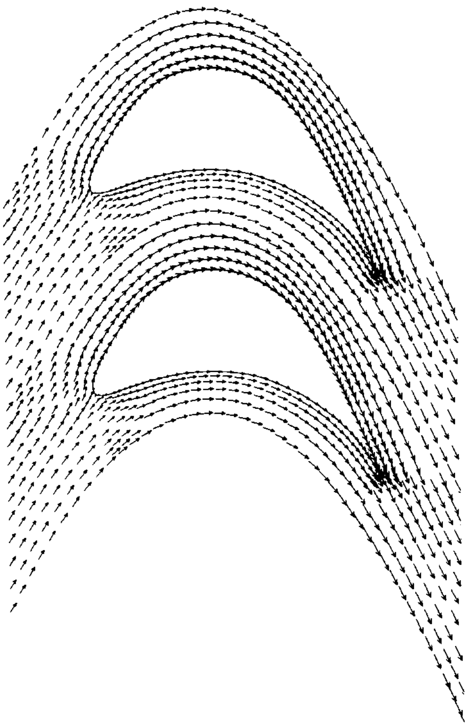

Fig. 6.12.

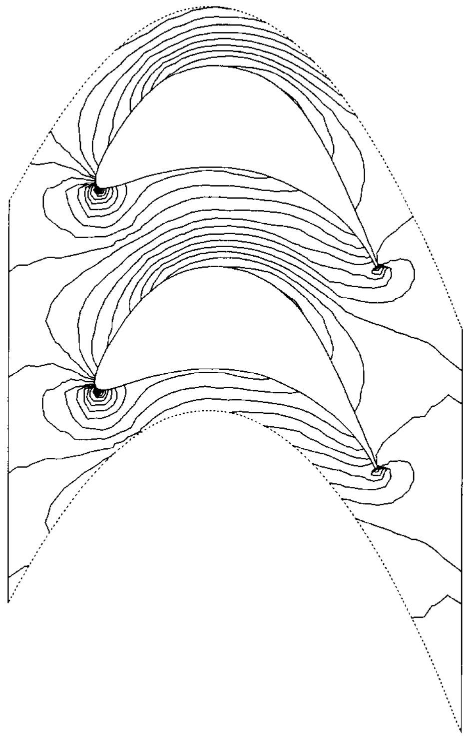

Fig. 6.13.

for the decrease of the residual in (3.17c) represented by $\| \bar{z}^{k+1} - \bar{z}^k \| / \| \bar{z}^{k+1} \|$ under 10^{-4} in dependence of ν.

In Fig. 6.10 we see the comparison of the incompressible and compressible velocity distribution along the profile in dependence of the length of the arc measured along the profile in the anti-clockwise direction starting from the trailing point. Figure 6.11 represents the distribution of Mach number along the profile in the compressible case. In Figs. 6.12 and 6.13 the velocity vectors and the isolines of the velocity, respectively, are plotted.

References

[1] O. Axelsson, Solution of linear systems of equations: Iterative methods, in: V.A. Barker, Ed., *Sparse Matrix Techniques*, Lecture Notes in Math. **572** (Springer, Berlin, 1977).

[2] I. Babuška, M. Práger and E. Vitásek, *Numerical Processes in Differential Equations* (SNTL, Prague/Wiley, New York, 1966).
[3] I. Černý, *Foundation of Analysis in the Complex Domain* (Academia, Prague/Ellis Horwood, Chichester, 1992).
[4] P.G. Ciarlet, *The Finite Element Method for Elliptic Problems*, Stud. Math. Appl. **4** (North-Holland, Amsterdam, 1978).
[5] P.G. Ciarlet and P.-A. Raviart, Maximum principle and uniform convergence for the finite element method, *Comput. Methods Appl. Mech. Engrg.* **2** (1) (1973) 17–31.
[6] J.F. Ciavaldini, M. Pogu and G. Tournemine, Existence and regularity of stream functions for subsonic flows past profiles with a sharp trailing edge, *Arch. Rational Mech. Anal.* **93** (1986) 1–14.
[7] M. Feistauer, Solution of elliptic problem with not fully specific Dirichlet boundary value conditions and its application in hydrodynamics, *Apl. Mat.* **24** (1979) 67–74.
[8] M. Feistauer, Mathematical study of three-dimensional axially symmetric stream fields of an ideal fluid, in: B. Brosowski and E. Martensen, Eds., *Methods and Techniques of Mathematical Physics*, Methoden Verfahren Math. Phys. **21** (Lang, Frankfurt am Main, 1981) 45–62.
[9] M. Feistauer, *Discrete Methods of the Solution of Differential Equations* (SPN, Prague, 1981, in Czech).
[10] M. Feistauer, Mathematical study of rotational incompressible non-viscous flows through multiply connected domains, *Apl. Mat.* **26** (1981) 345–364.
[11] M. Feistauer, Numerical solution of non-viscous channel flows, in: E. Meister, K. Nickel and J. Polášek, Eds., *Mathematical Methods in Fluid Mechanics*, Methoden Verfahren Math. Phys. **24** (Lang, Frankfurt am Main, 1982) 65–78.
[12] M. Feistauer, Subsonic irrotational flows in multiply connected domains, *Math. Methods Appl. Sci.* **4** (1982) 230–242.
[13] M. Feistauer, On irrotational flows through cascades of profiles in a layer of variable thickness, *Apl. Mat.* **29** (1984) 423–458.
[14] M. Feistauer, On the finite element approximation of a cascade flow problem, *Numer. Math.* **50** (1987) 655–684.
[15] M. Feistauer, Nonlinear elliptic problems with incomplete Dirichlet conditions and the stream function solution of subsonic rotational flows past profiles or cascades of profiles, *Apl. Mat.* **34** (4) (1989) 318–339.
[16] M. Feistauer, J. Felcman and Z. Vlášek, Finite element solution of flows through cascades of profiles in a layer of variable thickness, *Apl. Mat.* **31** (1986) 309–339.
[17] M. Feistauer and M. Huněk, Solution of an axially symmetric flow by the finite element method, *Stroj. Čas.* **34** (1983) 607–621 (in Czech).
[18] M. Feistauer and J. Římánek, Subsonic irrotational flow of compressible fluid in axially symmetric channels, *Apl. Mat.* **20** (1975) 266–279.
[19] J. Felcman, Numerical solution of flows in cascades of blades by the finite element method, Ph.D. Thesis, Charles Univ., Prague, 1986 (in Czech).
[20] A. Kufner, O. John and S. Fučík, *Function Spaces* (Academia, Prague, 1977).
[21] K.W. Morton, Finite element methods for non-self-adjoint problems, in: P.R. Turner, Ed., *Topics in Numerical Analysis*, Lecture Notes in Math. **965** (Springer, Berlin, 1982) 113–148.
[22] J. Nečas, *Les Méthodes Directes en Théorie des Équations Elliptiques* (Academia, Prague/Masson, Paris, 1967).
[23] M. Rokyta, Numerical solution of strongly nonlinear elliptic problems, Diploma Thesis, Charles Univ., Prague, 1985 (in Czech).
[24] R.S. Varga, *Matrix Iterative Analysis* (Prentice-Hall, Englewood Cliffs, NJ, 1962).